

5

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2003-513564

(P2003-513564A)

(43) 公表日 平成15年4月8日(2003.4.8)

(51) Int.Cl. ⁷	識別記号	F I	テーマコード* (参考)
H 0 4 N 5/91		H 0 4 N 5/91	Z 5 C 0 5 3
7/24		7/13	Z 5 C 0 5 9

審査請求 未請求 予備審査請求 未請求(全 54 頁)

(21) 出願番号 特願2001-534885(P2001-534885)
 (86) (22) 出願日 平成12年10月13日(2000.10.13)
 (85) 翻訳文提出日 平成13年7月4日(2001.7.4)
 (86) 国際出願番号 PCT/EP00/10161
 (87) 国際公開番号 WO01/033863
 (87) 国際公開日 平成13年5月10日(2001.5.10)
 (31) 優先権主張番号 09/433, 258
 (32) 優先日 平成11年11月4日(1999.11.4)
 (33) 優先権主張国 米国 (US)
 (81) 指定国 EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, I T, LU, MC, NL, PT, SE), JP, KR

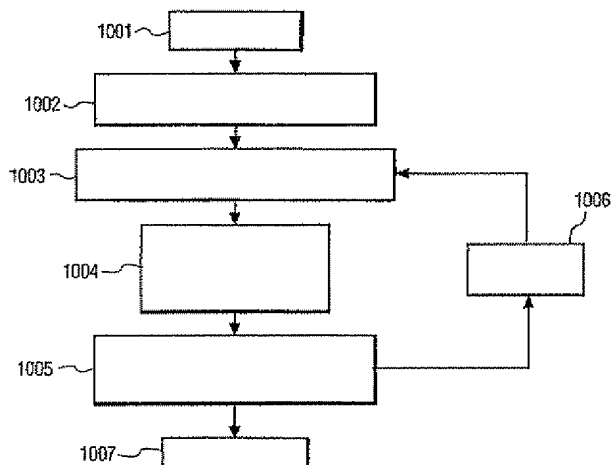
(71) 出願人 コーニンクレッカ フィリップス エレクトロニクス エヌ ヴィ
 Koninklijke Philips Electronics N. V.
 オランダ国 5621 ペーアー アイन्दーフエン フルーネヴァウツウェッハ 1
 Groenewoudseweg 1,
 5621 BA Eindhoven, The Netherlands
 (72) 発明者 マッギー, トマス
 オランダ国, 5656 アーアー アイन्दーフエン, プロフ・ホルストラーン 6
 (74) 代理人 弁理士 伊東 忠彦

最終頁に続く

(54) 【発明の名称】 動的閾値を使用する視覚的索引付けシステムの重要シーン検出及びフレームフィルタリング

(57) 【要約】

ビデオ索引付けシステムは、ソースビデオの内容を解析し、選択された画像を用いて内容の視覚的テーブルを作成する。重要シーン検出用システムは、あるシーンから別のシーンの間のビデオカットと、DCT係数及びマクロブロックに基づく静止シーンを検出する。主要フレームフィルタリング処理は、たとえば、単色のフレーム、又は、プライマリ・フォーカスと同一の対象を有するフレームのような余り望ましくないフレームを除去する。コマーシャルも検出され、コマーシャルのフレームは削除される。重要シーン及び静止シーンは、ビデオのカテゴリに基づいて設定された閾値に基づいて検出される。



【特許請求の範囲】

【請求項1】 ソースビデオの重要シーンを検出し、キーフレームを選択するシステムであって、

DCT係数のブロック及びフレームを含むマクロブロックとしてソースビデオを受信する受信機と、

k がフレームの幅方向のマクロブロック番号を表し、 $1 \leq k \leq \text{フレーム幅} / 16$ なる数であり、

j がフレームの高さ方向のマクロブロック番号を表し、 $1 \leq j \leq \text{フレーム高さ} / 16$ なる数であり、

i がマクロブロック内のブロック番号を表し、 $1 \leq i \leq \text{マクロブロック内のブロック数}$ であり、

$DCT1_{k,j}$ が前フレームに対する指定されたマクロブロックのDCT係数であり、 $DCT2_{k,j}$ が現在フレームに対する指定されたマクロブロックのDCT係数であり、

ABS が絶対値関数であり、

マクロブロック内のブロックのタイプ毎に、

$$SUM[i] = \sum_{k,j} ABS(DCT1_{k,j}[i] - DCT2_{k,j}[i])$$

に基づいて和 $SUM[i]$ を計算する加算器と、

ソースビデオのカテゴリに基づいて第1の閾値及び第2の閾値を調節する閾値調節器と、

各和 $SUM[i]$ を第1の閾値及び第2の閾値と比較し、和 $SUM[i]$ が第1の閾値よりも大きい場合に、現在フレームをキーフレームとしてフレームメモリに保存し、和 $SUM[i]$ が第2の閾値よりも小さい場合に、静止シーンカウンタをインクリメントし、前フレームを一時メモリに保存し、それ以外の場合に、静止シーンカウンタをリセットする第1の比較器と、

静止シーンカウンタを所定の数と比較し、一時メモリに保存された直前ビデオフレームをキーフレームとしてフレームメモリに転送する第2の比較器と、を有するシステム。

【請求項2】 ソースビデオの重要シーンを検出し、キーフレームを選択するシステムであって、

DCT係数のブロック及びフレームを含むマクロブロックとしてソースビデオを受信する受信機と、

k がフレームの幅方向のマクロブロック番号を表し、 $1 \leq k \leq \text{フレーム幅} / 16$ なる数であり、

j がフレームの高さ方向のマクロブロック番号を表し、 $1 \leq j \leq \text{フレーム高さ} / 16$ なる数であり、

i がマクロブロック内のブロック番号を表し、 $1 \leq i \leq \text{マクロブロック内のブロック数}$ であり、

$DCT1_{k,j,i}$ が前フレームに対する指定されたマクロブロック及びブロックのDCT係数であり、 $DCT2_{k,j,i}$ が現在フレームに対する指定されたマクロブロック及びブロックのDCT係数であり、

ABS が絶対値関数であり、

各マクロブロックに対し、

$$SUM = \sum_{k,j} \sum_i (DCT1_{k,j,i} - DCT2_{k,j,i})^2 / ABS(DCT2_{k,j,i})$$

に基づいて和 SUM を計算する加算器と、

ソースビデオのカテゴリに基づいて第1の閾値及び第2の閾値を調節する閾値調節器と、

各和 SUM を第1の閾値及び第2の閾値と比較し、和 SUM が第1の閾値よりも大きい場合に、現在フレームをキーフレームとしてフレームメモリに保存し、和 SUM が第2の閾値よりも小さい場合に、静止シーンカウンタをインクリメントし、前フレームを一時メモリに保存し、それ以外の場合に、静止シーンカウンタをリセットする第1の比較器と、

静止シーンカウンタを所定の数と比較し、一時メモリに保存された直前ビデオフレームをキーフレームとしてフレームメモリに転送する第2の比較器と、
を有するシステム。

【請求項3】 ソースビデオの重要シーンを検出し、キーフレームを選択す

るシステムであって、

DCT係数のブロック及びフレームを含むマクロブロックとしてソースビデオを受信する受信機と、

k がフレームの幅方向のマクロブロック番号を表し、 $1 \leq k \leq \text{フレーム幅} / 16$ なる数であり、

j がフレームの高さ方向のマクロブロック番号を表し、 $1 \leq j \leq \text{フレーム高さ} / 16$ なる数であり、

i がマクロブロック内のブロック番号を表し、 $1 \leq i \leq \text{マクロブロック内のブロック数}$ であり、

$DCT1_{k,j}$ が前フレームに対する指定されたマクロブロックのDCT係数であり、 $DCT2_{k,j}$ が現在フレームに対する指定されたマクロブロックのDCT係数であり、

ABS が絶対値関数であり、

マクロブロック内のブロックのタイプ毎に、

$$SUM[i] = \sum_{k,j} (DCT1_{k,j}[i] - DCT2_{k,j}[i])^2 / ABS(DCT2_{k,j})$$

に基づいて和 $SUM[i]$ を計算する加算器と、

ソースビデオのカテゴリに基づいて第1の閾値及び第2の閾値を調節する閾値調節器と、

各和 $SUM[i]$ を第1の閾値及び第2の閾値と比較し、和 $SUM[i]$ が第1の閾値よりも大きい場合に、現在フレームをキーフレームとしてフレームメモリに保存し、和 $SUM[i]$ が第2の閾値よりも小さい場合に、静止シーンカウンタをインクリメントし、前フレームを一時メモリに保存し、それ以外の場合に、静止シーンカウンタをリセットする第1の比較器と、

静止シーンカウンタを所定の数と比較し、一時メモリに保存された直前ビデオフレームをキーフレームとしてフレームメモリに転送する第2の比較器と、
を有するシステム。

【請求項4】 ソースビデオの重要シーンを検出し、キーフレームを選択するシステムであって、

DCT係数のブロック及びフレームを含むマクロブロックとしてソースビデオを受信する受信機と、

k がフレームの幅方向のマクロブロック番号を表し、 $1 \leq k \leq \text{フレーム幅} / 16$ なる数であり、

j がフレームの高さ方向のマクロブロック番号を表し、 $1 \leq j \leq \text{フレーム高さ} / 16$ なる数であり、

i がマクロブロック内のブロック番号を表し、 $1 \leq i \leq \text{マクロブロック内のブロック数}$ であり、

$MB1_{k,j}$ が前ビデオフレームのマクロブロックを表し、 $MB2_{k,j}$ が現在ビデオフレームのマクロブロックを表し、

$Mbdiff$ が、二つのマクロブロック間で相異なるブロック数を定め、この差がある閾値よりも大きい場合に第1の値を出力し、さもなければ、第2の値を出力する関数であり、

マクロブロック内のブロックのタイプ毎に、

$SUM = \sum_{k,j} Mbdiff (MB1[i]_{k,j} - MB2[i]_{k,j})$ に基づいて和 SUM を計算する加算器と、

を有し、

上記加算器は、 $DCT1_{k,j}$ 及び $DCT2_{k,j}$ が指定されたマクロブロックのDCT係数を表すとき、

各ブロックに対し

$subsum[i]_{j,k} = ABS(DCT1[i]_{j,k} - DCT2[i]_{j,k})$

に従って部分和 $subsum$ を計算する部分和計算器を含み、

上記システムは、

各部分和を第1の所定の部分和閾値と比較し、各部分和が第1の所定の部分和閾値よりも大きい場合に、第1のブロックカウンタをインクリメントし、各部分和を第2の所定の部分和閾値と比較し、各部分和が第2の所定の部分和閾値よりも小さい場合に、第2のブロックカウンタをインクリメントする部分和比較器と、

第1のブロックカウンタを第1の所定のブロック閾値と比較し、第1のブロックカウンタが第1の所定のブロック閾値よりも大きい場合に、第1のマクロブロックカウンタをインクリメントし、次のマクロブロックの解析のため第1のブロックカウンタをリセットし、第2のブロックカウンタを第2の所定のブロック閾値と比較し、第2のブロックカウンタが第2の所定のブロック閾値よりも大きい場合に、第2のマクロブロックカウンタをインクリメントし、次のマクロブロックの解析のため第2のブロックカウンタをリセットするブロック比較器と、

第1のマクロブロックカウンタを第1の所定のフレーム閾値と比較し、第1のマクロブロックカウンタが第1の所定のフレーム閾値よりも大きい場合に、現在フレームをフレームメモリに保存し、次のフレームの解析のため第1のマクロブロックカウンタをリセットし、第2のマクロブロックカウンタを第2の所定のフレーム閾値と比較し、第2のマクロブロックカウンタが第2の所定のフレーム閾値よりも大きい場合に、フレームカウンタをセットし、第2のマクロブロックカウンタをリセットするマクロブロック比較器と、

フレームカウンタを所定のシーン閾値と比較し、フレームカウンタが所定のシーン閾値よりも大きい場合に、前フレームをキーフレームとしてフレームメモリに保存するフレーム比較器と、

所定の部分和閾値、所定のブロック閾値、及び、所定のブロック閾値の内の少なくとも一つをソースビデオのカテゴリに基づいて調節する閾値調節器と、を更に有する、システム。

【請求項5】 ビデオソースからのフレームをフィルタリングするシステムであって、

フレームをフレームメモリに受信し、DCT係数のブロックからなるマクロブロックを一時メモリに受信する受信機と、

少なくとも一つの署名閾値に基づいてブロック毎にブロック署名を作成するブロック署名作成器と、

各フレームの領域と対応したサイズを検出する領域拡張器と、

各領域の重心を判定する重心判定器と、

領域内のブロックのうちの一つのブロック署名、対応したサイズ及び重心を各

領域に割り当てる領域署名作成器と、

領域署名に基づいて第1のフレームに対しフレーム署名を作成するフレーム署名作成器と、

フレーム差を判定するため第1のフレームの少なくとも一つの特徴を第2のフレームの少なくとも一つの特徴と比較し、フレーム差を少なくとも一つの閾値と比較し、フレーム差が閾値よりも大きい小さいかに応じてフレームを保存又は廃棄するフレーム比較器と、

を有し、

署名閾値と閾値のうちの少なくとも一つは、ビデオソースから受信されたビデオのカテゴリに基づいている、システム。

【請求項6】 各ブロックは、DC値を含むDCT係数により構成され、

ブロック署名は、対応したDC値が収まる区間を表すDC署名により構成される、

請求項5記載のシステム。

【請求項7】 単色フレームを検出し、同じDC署名を有するブロックの数を計数し、計数されたブロックの数が、到来ビデオのカテゴリに基づいて予め定められたカウントよりも大きい場合に、フレームを除去する単色フィルタを更に有する、

請求項6記載のシステム。

【請求項8】 到来ビデオのカテゴリに基づいて予め定められた規準を満たさない領域を除く領域フィルタを更に有する、請求項5記載のシステム。

【請求項9】 各ブロックは、DC値を含むDCT係数により構成され、

ブロック署名は、対応したAC値が所定の値を超えるかどうかを表現するように予め定められたビット数のビットのAC署名により構成され、

予め定められたビット数と所定の値のうちの少なくとも一方は到来ビデオのカテゴリに基づいている、

請求項5記載のシステム。

【請求項10】 フレーム内の各領域の対応したサイズを、到来ビデオのカ

テゴリーに基づく所定の数と比較し、対応したサイズが所定の数よりも大きい場合に、フレームメモリからフレームを廃棄する単色フィルタを更に有する、請求項5記載のシステム。

【請求項11】 ビデオのフレームを受信する受信機と、
ビデオの特定のカテゴリーに対する閾値を受け取る閾値メモリと、
第1のフレームと後続のフレームの間の差を判定するため第1のフレームを後続のフレームと比較し、差を閾値と比較し、差が閾値よりも大きい場合に、二つのフレームの間にシーン変化が生じたとみなし、差が閾値よりも小さい場合に、シーン変化が生じていないとみなす比較器と、
を有するビデオシステム。

【請求項12】 差が複数のフレームに対する閾値よりも小さい場合に、シーンは静止シーンであるとみなされる、請求項11記載のビデオシステム。

【請求項13】 ビデオのカテゴリーを検出するカテゴリー検出器と、
検出されたカテゴリーに基づいて閾値を設定する閾値設定器と、
を有する、請求項11又は12記載のビデオシステムで使用するための閾値設定装置。

【請求項14】 検出されたシーン変化の数が所定の数を超えると、閾値は自動的に変更され、検出されたシーン変化の数が所定の数に満たないとき、閾値は自動的に変更される、請求項11又は12記載のビデオシステム。

【請求項15】 各シーン変化からのキーフレームを受け取り、
冗長なキーフレーム、類似したキーフレーム、及び、単色のキーフレームのうちの少なくとも一つを除去する、
請求項11又は12記載のビデオシステムで使用するためのフィルタリングシステム。

【発明の詳細な説明】**【0001】**

本発明は、ソースビデオの重要シーンを検出し、検出された重要シーンを個々に表現するキーフレームを選択する装置に関する。本発明は、さらに、選択されたキーフレームをフィルタ処理し、残りのキーフレームに基づいて内容の視覚的索引又は視覚的テーブルを作成する。

【0002】

ユーザーは、ホームビデオ、テレビジョン番組、映画、コンサート、スポーツイベントなどをテープに録画し、後で視聴し、或いは、繰り返し視聴する。屢々、ビデオは、種々の内容を有し、長くなる。しかし、ユーザーは、テープに何が録画されているかを書き留めることがなく、テープ、DVD若しくはその他の媒体に何を録画したかを記憶していないし、特定のシーン、映画、若しくは、イベントがテープのどの場所に録画されているかを記憶していない。そのため、ユーザーは、テープに何が録画されているかを思い出すためにじっとしてテープ全体を視聴しなければならない。

【0003】

ビデオ内容解析は、録画された素材の内容を記述する情報を取り出すため自動的又は半自動的な方法を使用する。ビデオ内容索引付け及び解析は、ビデオ内の視覚的キューから構造及び意味を抽出する。一般的に、ビデオクリップは、TV番組若しくはホームビデオから選ばれる。

【0004】

米国特許出願第08/867140号には、ビデオ中のシーン変化又はカットを検出する方法及び装置が記載され、本願はこの米国特許出願の一部継続出願である。検出されたカットの間の少なくとも一つのフレームは、ビデオ索引付けを作成するためキーフレームとして選択される。シーン変化を検出するため、最初のフレームが選択され、次に、後続のフレームは最初のフレームと比較され、二つのフレーム間の内容差を表現する差が計算される。この差の計算の結果は、ビデオの全てのカテゴリーに対して使用される汎用的な閾値若しくは閾値と比較される。

【0005】

上述の米国特許出願第08/867140号によれば、全てのタイプのビデオに対し最適である汎用的な閾値が選択される。この特許出願における問題は、アクション映画のような激しい展開を含むビデオの視覚的索引は非常に大きく、一方、ニュースのように展開の少ないビデオの視覚的索引は非常に小さいことである。なぜならば、展開の激しい映画の場合、物体がシーン全体を移動するとき、二つの連続したフレーム間の内容の差が大きいからである。このような場合に、内容差を汎用的な閾値と比較することにより、二つのフレームが同一シーン内に存在する場合でもカットが検出される。より多くのカット若しくはシーン変化が認知される場合、より多くのキーフレームが存在し、その逆に、より多くのキーフレームが存在する場合、より多くのカット若しくはシーン変化が認知される。これに応じて、アクション映画は、その映画を表現するため、最終的に非常に多数のキーフレームを含む。

【0006】

したがって、本発明の目的は、予め録画されたビデオソース、又は、録画中のビデオソースに対する視覚的索引を作成するシステムであって、ビデオのカテゴリに基づいてビデオから選択されたキーフレームの数を変化させることにより、重要なキーフレームを選択する際に使用される非常に精密なシステムを提供することである。

【0007】

本発明は、さらに、ニュース、続きもの、気象、スポーツ、又は、任意のタイプのホーム録画されたビデオ、放送素材、予め録画された素材などのソースビデオのための視覚的内容抽出を支援するビデオ解析システムを提供する。

【0008】

本発明は、さらに、ビデオカット検出、静止シーン検出、及び、キーフレームフィルタリング処理用の新しい装置であって、ビデオの二つのフレームを比較し、フレーム間の差が、ビデオのカテゴリに依存したある種の閾値を上回るか、若しくは、下回るかを決定することにより、視覚的索引により有用性のある視覚的画像を与えることができる装置を提供することである。差が選択された閾値を上回る場合、ビデオ中のこの点はシーンの「カット」であると判定される。シー

ン・カット内のフレームは、ビデオ索引用のキーフレームを選択するためフィルタ処理され得る。差が多数のフレームに対する静止シーン閾値に達しないとき、静的シーケンスであると判定される。

【0009】

本発明は、DCT係数を使用する計算と、たとえば、アクション、ニュース、音楽のようなビデオのカテゴリー、又は、ビデオの一部分のカテゴリーに応じて変化する種々の閾値との比較とに基づいて、重要シーンを検出し、ソースビデオのキーフレームを選択する。

【0010】

したがって、本発明の目的は、電子番組ガイドによって与えられるビデオカテゴリーに基づいてこれらの閾値を計算すること、或いは、閾値を計算する代わりに、電子番組ガイド自体がこれらの閾値を与えることである。

【0011】

さらに、本発明の目的は、これらの閾値を手動で入力することである。

【0012】

本発明のさらに別の目的は、符号化されたビデオに閾値を設けることである。

【0013】

本発明と、本発明の効果と、本発明を使用することにより達成される具体的目的とをよりよく理解するため、以下では、添付図面を参照して、本発明の好ましい実施例を説明する。

【0014】

[視覚的索引付けシステム]

最初に、視覚的索引付けシステムの基本動作について説明する。

【0015】

視覚的索引は、既存のテープ（若しくは、ファイル、DVD、ディスクなど）に作成され、又は、新しいテープへの記録時に作成される。既存テープ及び新テープの両方の場合に、（利用可能であれば）テープ上の選択された領域の所定の部分、本例の場合には、利用の容易性のため先頭部分は、視覚的索引を作成することが許容され、或いは、索引は、ハードディスクなどの全く異なる媒体に格納

してもよい。本例の場合、30秒の「ブランク（空白）」又は上書き可能なテープが望ましい。ファイルの場合、視覚的索引のため選択された領域はファイル内のどのような場所でも置くことができ、システムによって自動的に、又は、ユーザーによって手動で選択されて確保される。

【0016】

視覚的索引には、視覚的画像、オーディオ、テキスト、又は、コレラの組み合わせが含まれる。本例の場合には、視覚的画像とテキストが設けられる。視覚的索引を作成し、使用するため、ビデオ内容索引付け処理が実行される。

【0017】

ビデオ内容索引付け処理には、保管と検索の二つの段階が存在する。保管処理中に、ビデオ内容が解析され、ビデオ解析中に視覚的索引が作成される。ビデオ解析処理において、自動重要シーン検出及びキーフレーム選択が行なわれる。重要シーン検出は、シーン変化、すなわち、カットを識別し（ビデオカット検出、又は、セグメンテーション検出）、静止シーンを識別する（静止シーン検出）処理である。シーン毎に、特定の代表的なフレーム、すなわち、キーフレームが抽出される。キーフレームフィルタリング及び選択処理は、キーフレームから選択的に選ばれた視覚的索引を作成するため、ビデオテープのようなソースビデオの各キーフレームに適用される。一例として、ソーステープを挙げているが、勿論、ソースビデオは、ファイル、ディスク、DVD、その他の記憶手段から、又は、（たとえば、ホームビデオを録画中に）伝送ソースから直接的に得ることができる。

【0018】

ビデオテープ索引付けの場合に、索引は、一般的に、ソーステープ、CDなどに記憶される。MPEG1、MPEG2、MPEG4、及び、モーションJPEGファイル、ビデオCD、DVD若しくはその他の記憶装置からのその他のビデオファイル、又は、放送ストリームからのビデオファイルをビデオ索引付けする場合、索引は、ハードディスク、若しくは、その他の記憶媒体に記憶される。

【0019】

図1には、ソーステープが予め記録されたソーステープであり、オーディオ及

び／又はテキストを含む場合のビデオ保管処理は示されている。しかし、MPEGファイルのような視覚的情報が予め格納された他の記憶装置の場合にも同様の処理が続けられる。この処理において、視覚的索引は、ソースビデオに基づいて作成される。ユーザーが記録しようとしているソーステープの場合、第2の処理は、記録と同時に視覚的索引を作成する。

【0020】

図1には、（記録済みのソーステープの場合の）ビデオテープ用の第1の処理の一例が示されている。ステップ101において、ソーステープは、必要であれば、VCRのような再生／録画装置で巻戻される。ステップ102において、ソースビデオが再生される。ソースビデオからの信号は、テレビジョン、VCR、或いは、その他の処理装置によって受信される。ステップ103において、処理装置内のメディアプロセッサ、又は、外部プロセッサは、ビデオ信号を受信し、ビデオ信号を、画素データを表現するフレームの形式にフォーマットする（フレーム獲得）。

【0021】

ステップ104において、ホストプロセッサは、各フレームをブロックに分離し、DCT（離散コサイン変換）係数を作成するためブロック及びブロックの関連データを変換し、重要シーン検出及びキーフレーム選択を行い、キーフレームをデータ構造としてメモリ、ディスク、若しくは、その他の記憶媒体に格納する。ステップ105において、ソーステープは先頭まで巻き戻され、ソーステープは情報を記録するためセットされる。ステップ107において、データ構造がメモリからソーステープへ転送され、視覚的索引が生成される。テープは、次に、巻き戻され、視覚的索引が参照される。

【0022】

ユーザーが録画中にテープに視覚的索引を作成したい場合、上記の処理は僅かに変更される。ステップ101及び102の代わりに、図1にステップ112で示されるように、ステップ103のフレーム獲得処理が、ビデオ（映画など）の録画と同時に行なわれる。

【0023】

さらに、テープ又はファイルが、一度に完全に録画されない場合、部分的に作成されたビデオ索引がテープ、ファイルなどに保存され、或いは、後で追加するためテープメモリに保存される。

【0024】

ステップ103及び104は、図2A及び2Bにより詳細に示されている。ビデオは、アナログ（連続的データ）形式、又は、デジタル（離散的データ）形式のいずれかの形式である。本例は、デジタル域で動作し、処理の際にデジタル形式を使用する。ソースビデオ又はビデオ信号は、非常に高いレート（本例の場合には、毎秒30フレーム）で表示される個別の画像、若しくは、ビデオフレームの系列であり、表示された画像のシーケンスは、連続したピクチャーストリームとして現れる。これらのビデオフレームは、非圧縮形式（NTSC若しくは未加工のビデオ）、又は、MPEG、MPEG2、MPEG4、モーションJPEGなどのような形式で圧縮されたデータである。

【0025】

非圧縮ビデオ内の情報は、最初に、Intel[®] Smart Video Recorder IIIで得られるようなフレーム獲得技術を使用して、メディアプロセッサ202でフレームにセグメント化される。図3に示された例の場合に、他のフレームサイズを使用してもかまわないが、フレーム302は、一つのテレビジョン画像、ビデオ画像、若しくは、その他の視覚的画像を表現し、352×240画素を含む。

【0026】

各フレーム302は、本例の場合には、ホストプロセッサ210（図2A）で、8×8画素からなるブロック304に分割される。これらのブロック304と、普及している放送標準CCIR-601とを使用して、マクロブロック作成器206（図2A）は、ルミナンスブロックと、クロミナンスブロックを作成するための平均カラー情報とを作成する。ルミナンスブロック及びクロミナンスブロックは、マクロブロック308を形成する。本例の場合、4:2:0が使用されるが、当業者は、4:1:1や4:2:2のようなその他のフォーマットを簡単に使用することができる。4:2:0フォーマットの場合、マクロブロック308は、4個のルミナンスブロックY1、Y2、Y3及びY4と、2個のクロミナ

ンスブロックC_r及びC_bの併せて6個のブロックを含み、マクロブロック内の各ブロックは、8×8画素により構成される。

【0027】

ビデオ信号は、モーションJ P E G（ジョイント・フォトグラフィック・エクスパート・グループ）及びM P E G（モーション・ピクチャー・エクスパート・グループ）のような圧縮標準を使用して圧縮信号を表現してもよい。図2Bに示されるように、信号がM P E G圧縮信号若しくはその他の圧縮信号である場合、M P E G信号は、フレームパーサー205によって、フレーム若しくはビットストリーム解析技術を使用して分割される。次に、フレームは、メディアプロセッサ203内のエントロピー復号器214と、テーブル指定器216とに送られる。エントロピー復号器214は、たとえば、ハフマン復号化方式又はその他の復号化方式を使用し、テーブル指定器216からのデータを用いてM P E G信号を復号化する。

【0028】

復号化信号は、逆量子化器218へ送られ、逆量子化器218はテーブル指定器216からのデータを用いて復号化信号を逆量子化する。これらのステップ（ステップ214－218）は、メディアプロセッサ203で行なわれるものとして説明しているが、使用される装置に応じて、メディアプロセッサ203、ホストプロセッサ211、或いは、別の外部装置で行なうことも可能である。

【0029】

或いは、システム（たとえば、メディアプロセッサ）に、処理の種々の段階へアクセスすることができる符号化機能が備わっている場合、D C T係数は、ホストプロセッサへそのまま配布してもよい。上述の全てのアプローチにおいて、処理は実時間で行なわれる。

【0030】

図1のステップ104において、たとえば、Intel^(R) PentiumTM チップ若しくはその他のマルチプロセッサ、又は、Phillips^(R) TrimediaTM チップ若しくはその他のマルチメディアプロセッサ、コンピュータ、拡張V C R、録画／再生装置、テレビジョン、或いは、その他のプロセッサにより実現されるホストプロ

セッサ210は、重要シーン検出、キーフレーム選択、並びに、たとえば、ハードディスク、ファイル、テープ、DVD若しくはその他の記録媒体のような索引メモリにおけるデータ構造の構築及び記憶を実行する。

【0031】

[重要シーン検出]

自動重要シーン検出のため、本発明は、ビデオのシーンが変化したとき、或いは、静止シーンが出現したときの検出を試みる。シーンは、一つ以上の関連した画像を表現する。重要シーン検出において、二つの連続したフレームが比較される。これらのフレームが著しく相異していると判定された場合、二つのフレームの間でシーン変化が発生したと判定される。二つのフレームが非常に類似していると判定された場合、静止シーンが発生したかどうかの判定が行なわれる。ニュース放送のように静止タイプのシーンの場合、二つのフレーム間の僅かな変化がシーン変化を示すことがある。アクション映画の場合、二つのフレーム間の同程度の変化はシーン変化ではなく、シーン内の物体の移動に過ぎない場合がある。

【0032】

シーン毎に、シーンを表現するため一つ以上のキーフレームが抽出される。典型的に、現時点での理論は、シーン内の最初のビデオ（視覚的）フレームの使用を提案する。しかし、殆どの場合に、シーン内の主被写体又はイベントは、カメラズーム若しくはカメラパンの後に現れる。さらに、現時点での理論は、典型的に、ある時間の長さに亘って一定状態を保つシーン（静止シーン）を検出しない。しかし、そのシーン上で経過した時間の長さに基づいて、ビデオ撮影者、監督などの観点から、このシーンが重要シーンであるかもしれない。

【0033】

本例における各方法は、DCT（離散コサイン変換）係数の比較を使用する。ヒストグラム差、ウェーブレット、画素に基づく比較、エッジ、フラクタル、その他のDCT法、或いは、フレームを比較するため閾値を使用するその他の方法などの他の方法を使用してもよい。第一に、各受信フレーム302は、マクロブロック308を作成するため、ホストプロセッサ210で個別に処理される。ホストプロセッサ210は、DCT係数を抽出し、DCT係数からなる6個のサイ

ズ8×8のブロックを作成するため、離散コサイン変換220を使用して、空間情報を含む各マクロブロック308を処理する。受信されたビデオ信号がMPEGのように圧縮ビデオフォーマットである場合、DCT係数は、逆量子化後に抽出され、離散コサイン変換器によって処理する必要はない。さらに、上述の通り、DCT係数は、使用された装置に応じて自動的に抽出される。

【0034】

DCT変換器は、各ブロック440（図4）、Y1、Y2、Y3、Y4、Cr及びCbにDCT係数値を与える。この標準にしたがって、各ブロックの左上隅は、DC情報（DC値）を収容し、残りのDCT係数は、AC情報（AC値）を収容する。AC値は、図4に部分的に示されるように、DC値の右側からDCT値の真下にあるDCT係数へジグザグ的な順序で周波数が増大する。

【0035】

本発明は、幾つかの異なる重要シーン検出方法を使用する。これらの方法は、すべて、対応したブロックに対するDCT係数を使用する。ホストプロセッサ210は、重要シーンプロセッサ230（図2A）内で、以下の方法のうちの少なくとも一つの方法を用いて各マクロブロックを処理する。

【0036】

以下の方法で、処理は、できるだけ迅速に結果を生成し、効率に著しい損失を伴わない処理に制限するため、DC値に制限される。しかし、当業者が各マクロブロックに対し全てのDCT係数を処理し得ることは明白である。これらのステップを使用して各ブロックを巡回することにより、全てのDCT係数が解析されるが、処理に要する時間に影響を与える。

【0037】

〔方法1〕

$$SUM[i] = \sum_{k,j} ABS(DCT1_{k,j}[i] - DCT2_{k,j}[i])$$

但し、kは、フレームの幅方向のマクロブロック番号を表し、 $1 \leq k \leq \text{フレーム幅} / 16$ なる数であり、

jは、フレームの高さ方向のマクロブロック番号を表し、 $1 \leq j \leq \text{フレーム高さ}$

／16なる数であり、

i はマクロブロック内のブロック番号を表し、 $1 \leq i \leq$ マクロブロック内のブロック数であり、

図5に示されるように、 $DCT1_{k,j}$ は、前フレームに対する指定されたマクロブロックのDCT係数であり、 $DCT2_{k,j}$ は、現在フレームに対する指定されたマクロブロックのDCT係数であり、

ABS は絶対値関数である。

【0038】

本例の場合、 352×240 画素のフレーム、 $1 \leq k \leq 22$ 、 $1 \leq j \leq 15$ 、 $1 \leq i \leq 6$ である。この方法1及び以下の方法において、フレームのマクロブロック幅又はフレームのマクロブロック高さは偶数である。その理由は、フレームサイズが均等に分割できない場合、フレームサイズのスケールが処理中に調整されるからである。

【0039】

方法1は、マクロブロックの各マクロブロック（4個のルミナンスブロックと2個のクロミナンスブロック）の差分をとる。方法1において、現在ビデオフレームからの現在マクロブロック内の各ルミナンスブロック及び各クロミナンスブロックに対するDC値は、それぞれ、前ビデオフレーム内の対応したブロックに対する対応したDC値から減算される。別々の差分の和 $SUM[i]$ は、マクロブロック内のルミナンスブロック及びクロミナンスブロック毎に保持される。

【0040】

差分の和は、次のマクロブロックへ引き渡され、対応した差分（ $SUM[1]$ 、 $SUM[2]$ 、...、 $SUM[6]$ ）へ加算される。現在ビデオフレームの各マクロブロックを処理した後、合計された差分が、現在ビデオフレームの各ルミナンスブロックと各クロミナンスブロックに対し獲得される。6個の和 SUM は、それぞれ、各 SUM が集計されたブロックのタイプに固有の上側閾値及び下側閾値と比較される。この方法は、ブロックのタイプ毎に異なる閾値比較を行なう。

【0041】

SUM [i] が、本例では、

$$\text{thresh1}[i] = 0.3 * \text{ABS}(\sum_k, j \text{ DCT2}_{k,j}[i])$$
 である所定の閾値 (thresh1 [i]) よりも大きい場合、現在ビデオフレームは、後の処理のため、並びに、視覚的索引に使用する可能性があるためフレームメモリに保存される。フレームメモリは、本発明の場合のように、テープ、ディスク、或いは、本システムの外部若しくは内部に設けられたその他の記録媒体でもよい。

【0042】

SUM [i] が、

$$\text{thresh2}[i] = 0.02 * \text{ABS}(\sum_k, j \text{ DCT2}_{k,j}[i])$$
 である所定の閾値 (thresh2 [i]) よりも小さい場合、静止シーンである可能性を示すため静止シーンカウンタ (SSctr) を増加させる。前ビデオフレームは一時メモリに保存される。一時メモリは1フレームしか保存しないので、前ビデオフレームは、このとき一時メモリに保存されているビデオフレームを置き換える。カウンタが所定の数 (本例の場合には、30) に到達したとき、一時メモリに保存された直前ビデオフレームは、視覚的索引に使用される可能性があるためフレームメモリへ転送される。最後に続くフレームは、静止シーンを表現する可能性があるため保存される様子を説明したが、当業者であれば、当然に、この方法1及び以下の方法において、静止シーン候補の最初のフレームを保存し、使用し得る。

【0043】

SUM [i] が二つの閾値の間に取まるとき、SSctrは零にリセットされ、次の連続したフレームが比較される。

【0044】

本発明によれば、閾値は、カットレート、又は、ビデオ中のアクションの量に基づいて変更される。これに関しては、セクション：閾値判定において後述する。

【0045】

[方法2]

$$SUM = \sum_{k,j} \sum_i (DCT1_{k,j,i} - DCT2_{k,j,i})^2 / ABS(DCT2_{k,j,i})$$

但し、 k は、フレームの幅方向のマクロブロック番号を表し、 $1 \leq k \leq \text{フレーム幅} / 16$ なる数であり、

j は、フレームの高さ方向のマクロブロック番号を表し、 $1 \leq j \leq \text{フレーム高さ} / 16$ なる数であり、

i はマクロブロック内のブロック番号を表し、 $1 \leq i \leq \text{マクロブロック内のブロック数}$ であり、

$DCT1_{k,j,i}$ は、前フレームに対する指定されたマクロブロック及びブロックのDCT係数であり、 $DCT2_{k,j,i}$ は、現在フレームに対する指定されたマクロブロック及びブロックのDCT係数であり、

ABS は絶対値関数である。

【0046】

方法2は、方法1に対し、ブロックタイプを区別しない。その代わりに、方法2は、現在ビデオフレームと前ビデオフレームのマクロブロック間におけるDC差分の移動和を維持する。

【0047】

ブロック間の各差分は、平方され、現在ブロックのDCT値に対し正規化される。特に、現在ビデオフレームからのブロックのDCT値は、前ビデオフレーム内の対応したブロックの対応したDCTから減算される。この差は、次に、平方され、現在ビデオフレームの対応したDCT値で除算される。現在ビデオフレームのDCT値が零である場合、その比較に関する和は1にセットされる。フレームの各マクロブロックのブロック毎のDCT値に関する差は、全体として合計され、総和SUMが得られる。

【0048】

総和SUMは、次に、所定の閾値と比較される。SUMが、本例では、 $threshold = 0.3 * ABS(\sum_{k,j,i} DCT2_{k,j,i})$ である所定の閾値 ($threshold$) よりも大きい場合、現在ビデオフレームは、後の処理のため、並びに、視覚的索引に使用する可能性があるためフレームメ

メモリに保存される。

【0049】

SUMが、

$$t h r e s h 2 = 0.02 * A B S \left(\sum_{k,j,i} D C T 2_{k,j,i} \right)$$

である所定の閾値 (t h r e s h 2) よりも小さい場合、静止シーンである可能性を示すため静止シーンカウンタ (S S c t r) を増加させる。方法1と同様に、前ビデオフレームは、直前ビデオフレームだけを格納する一時メモリに保存される。S S c t rカウンタが所定の数 (本例の場合には、30) に到達したとき、一時メモリに保存された直前ビデオフレームは、視覚的索引に使用される可能性があるためフレームメモリへ転送される。SUMが二つの閾値の間に収まるとき、S S c t rは零にリセットされ、次の連続したフレームが比較される。

【0050】

本発明によれば、閾値は、カットレート、又は、ビデオ中のアクションの量に基づいて変更される。これに関しては、セクション：閾値判定において後述する。

【0051】

[方法3]

$$S U M [i] = \sum_{k,j} \left(D C T 1_{k,j} [i] - D C T 2_{k,j} [i] \right)^2 / A B S \left(D C T 2_{k,j} \right)$$

但し、kは、フレームの幅方向のマクロブロック番号を表し、 $1 \leq k \leq \text{フレーム幅} / 16$ なる数であり、

jは、フレームの高さ方向のマクロブロック番号を表し、 $1 \leq j \leq \text{フレーム高さ} / 16$ なる数であり、

iはマクロブロック内のブロック番号を表し、 $1 \leq i \leq \text{マクロブロック内のブロック数}$ であり、

D C T 1_{k,j}は、前フレームに対する指定されたマクロブロックのD C T係数であり、D C T 2_{k,j}は、現在フレームに対する指定されたマクロブロックのD C T係数であり、

A B Sは絶対値関数である。

【0052】

方法3は、方法1と同様に、マクロブロックの各ブロック（4個のルミナンスブロックと2個のクロミナンスブロック）の間の差分をとる。方法3の場合、現在ビデオフレームからの現在マクロブロックにおけるルミナンス及びクロミナンスの各ブロックに対するDC値は、それぞれ、前ビデオフレーム内の対応したブロックに対する対応したDC値から減算される。しかし、方法2と同様に、ブロック間の各差分は、平方され、現在ブロックのDCT値に関して正規化される。特に、現在ビデオフレームからのブロックのDCT値は、前ビデオフレーム内の対応したブロックの対応したDCT値から減算される。この差は、次に平方され、現在ビデオフレームの対応したDCT値によって除算される。現在ビデオフレームDCT値が零である場合、その比較に対する和は1にセットされる。

【0053】

マクロブロック毎のブロックの各タイプの各DCT値の差は、一つに加算され、ブロックのタイプ毎の総和SUM[i]が得られる。別々の差分の和SUM[i]は、マクロブロック内のルミナンスブロック及びクロミナンスブロック毎に保持される。差分の和は、次のマクロブロックへ引き渡され、対応した差分（SUM[1]、SUM[2]、...、SUM[6]）へ加算される。現在ビデオフレームの各マクロブロックを処理した後、加算された差分が、現在ビデオフレームの各ルミナンスブロックと各クロミナンスブロックに対し獲得される。6個の和SUMは、それぞれ、各SUMが集計されたブロックのタイプに固有の上側閾値及び下側閾値と比較される。この方法は、ブロックのタイプ毎に異なる閾値比較を行なう。

【0054】

SUM[i]が、所定の閾値（既に定義済みのthreshold[i]）よりも大きい場合、現在ビデオフレームは、後の処理のため、並びに、視覚的索引に使用する可能性があるためフレームメモリに保存される。

【0055】

SUM[i]が、所定の閾値（既に定義済みのthreshold2[i]）よりも小さい場合、静止シーンである可能性を示すため静止シーンカウンタ（SSct

r)を増加させる。前ビデオフレームは一時メモリに保存される。一時メモリは、本例の場合、直前ビデオフレームだけを保存する。カウンタが所定の数、たとえば、30に到達したとき、一時メモリに保存された直前ビデオフレームは、視覚的索引に使用される可能性があるためフレームメモリへ転送される。

【0056】

SUM[i]が二つの閾値の間に収まるとき、SSctrは零にリセットされ、次の連続したフレームが比較される。

【0057】

本発明によれば、閾値は、カットレート、又は、ビデオ中のアクションの量に基づいて変更される。これに関しては、セクション：閾値判定において後述する。

【0058】

[方法4]

方法1乃至方法3は、完全なビデオフレームに対し作用し、全体的に加算され、若しくは、成分毎に合計された全てのルミナンス及びクロミナンスに対するDCT値の差分若しくは差分の2乗を加算する。方法4は、マクロブロックレベルで動作し、処理の制限された効率的な結果を与える。

【0059】

$SUM = \sum_{k,j} Mbdiff(MB1[i]_{k,j} - MB2[i]_{k,j})$
 但し、kは、フレームの幅方向のマクロブロック番号を表し、 $1 \leq k \leq \text{フレーム幅} / 16$ なる数であり、
 jは、フレームの高さ方向のマクロブロック番号を表し、 $1 \leq j \leq \text{フレーム高さ} / 16$ なる数であり、
 iはマクロブロック内のブロック番号を表し、 $1 \leq i \leq \text{マクロブロック内のブロック数}$ であり、
 $MB1_{k,j}$ は前ビデオフレームのマクロブロックを表し、 $MB2_{k,j}$ は現在ビデオフレームのマクロブロックを表し、
 Mbdiffは、二つのマクロブロック間で相異なるブロック数を定め、この差がある閾値よりも大きい場合に第1の値を出力し、さもなければ、第2の値を出

力する関数である。

【0060】

特に、部分和 (subsum [1], subsum [2], . . . subsum [6]) は、特定のマクロブロックのブロック (Y1, Y2, Y3, Y4, Cr, Cb) 毎に、第1のマクロブロックの夫々のブロックを、第2のマクロブロックの対応した夫々のブロックと比較することによって決められ、部分和 subsum [i] が得られる。ここで、

$$\text{subsum}[i]_{j,k} = \text{ABS}(\text{DCT1}[i]_{j,k} - \text{DCT2}[i]_{j,k})$$

である。

【0061】

たとえば、現在フレームの第1のマクロブロックのCrのDC値は、前フレームの第1のマクロブロックのCrのDC値から減算され、subsum [Cr]_{1,1} が得られる。各subsum [i] は、本例の場合に、

$$\text{th1} = 0.3 * \text{subsum}[i]$$

である所定の閾値 (th1) と比較される。subsum [i] が、本例の場合に、第1の閾値 (th1) よりも大きい場合、ブロックカウンタ (Blctr) はインクリメントされ、subsum [i] が、

$$\text{th2} = 0.02 * \text{subsum}[i]$$

である第2の閾値 (th2) よりも小さい場合、ブロックカウンタ (B2ctr) がインクリメントされる。対応した各subsum [i] は、閾値 (th1 及び th2) と比較される。閾値は、定数でもよく、固定関数に基づく値でもよく、又は、ブロックのタイプに特有の関数若しくは定数に基づく値でもよい。

【0062】

マクロブロックの6個のブロックが処理された後、ブロックカウンタが解析される。ブロックカウンタBlctrが、本例の場合に、3である所定の閾値 (Blth) を超える場合、このマクロブロックは、前ビデオフレームの対応したマクロブロックとは相異するとみなされ、マクロブロックカウンタMBlctrをインクリメントする。Blctrカウンタをリセットし、次のマクロブロックを

解析する。

【0063】

ビデオフレームの全てのマクロブロックが処理されたとき、MB1ctrを所定のフレーム閾値と比較する。MB1ctrが、352×240型のフレーム（若しくは画像）を使用する本例の場合に、第1の所定のフレーム閾値（f1th）、たとえば、100を超えると、現在フレームをフレームメモリに保存し、MB1ctrをリセットする。

【0064】

マクロブロック内である個数のブロックが類似している場合、B2ctrは、所定の閾値（B2th）、たとえば、2を超える場合、マクロブロックは同一であるとみなされ、第2のマクロブロックカウンタMB2ctrをインクリメントする。B2ctrをリセットし、次のマクロブロックを解析する。フレームの全てのマクロブロックを解析した後、第2のマクロブロックカウンタが第2のフレーム閾値（f2th）、たとえば、250を超えると、ビデオフレームは、同一であるとみなされ、フレームカウンタ（Fctr）がセットされる。MB2ctrはリセットされ、次のフレームが解析される。Fctrが所定の閾値（本例の場合に30であるSSthresh）に達したとき、静止シーケンスが出現したとみなす。前ビデオフレームは、フレームメモリに保存され、静止シーケンスを表現する。この処理は、ビデオフレームが前ビデオフレームと相異していると判定されるか、或いは、新フレームが使い尽くされるまで続けられる。次に、Fctrをリセットし、次のビデオフレームを解析する。

【0065】

方法4並びに方法1乃至3においてフレームメモリに保存されるこれらのフレームは、キーフレームであるとみなされる。

【0066】

方法4は、差分の正規化された平方を使用することによっても実現される。特に、ブロック間の差分を使用するだけではなく、この差分を平方し、減算されたブロックで見つけられた値によって除算してもよい。シーンカット検出の閾値に対する感度は低くなる。

【0067】

後述のキーフレームフィルタリング処理は、重要シーン検出処理の下で各フレームが処理されるときに、又は、全てのフレームが処理された後に、行なわれる。さらに、上述の閾値は、検出の量を増減させるため容易に変更される。たとえば、値0は、0.02のような他の値に容易に変更することが可能であり、或いは、定数は、たとえば、効率を増減させるため変更され、SSC trは同じでなくてもよい。さらに、各閾値は、定数、固定関数、又は、解析されるブロックのタイプ若しくは場所に応じて変化する関数でもよい。

【0068】

本発明によれば、閾値は、カットレート、又は、ビデオ中のアクションの量に基づいて変更される。これに関しては、セクション：閾値判定において後述する。

【0069】

[閾値判定]

上述の通り、ビデオ中のシーンカットが出現する場所の検出は、二つのフレーム間でどの程度の変化が生じるかに基づいている。変化の量は、少なくとも一つの閾値に対して測定され、変化がこの閾値を超える場合、シーン変化、すなわち、カットの発生が判定される。アクション映画の場合に、二つのフレーム間に大きい差が生じるが、シーン変化は出現しないことが頻繁にある。たとえば、カーチェイスのシーン、或いは、物体が空中を飛行するときにこれに該当する。同様に静止タイプの番組の場合、たとえば、ニュース放送は、ニュースストーリーを示す小さいボックスがフレームの右上隅に設けられ、フレームの右上隅がニュースストーリー間で変化したとき、この小さい変化はシーン変化を表す。したがって、番組若しくは映画のカテゴリーに応じて閾値を変更する必要がある。図8Aには、シーン変化を判定するため動的閾値を使用するビデオ保管システムが示されている。

【0070】

図9は、映画のカテゴリーに基づいて閾値を設定し、二つのフレーム間で見つけられた差をこの閾値と比較する閾値設定器のフローチャートである。

【0071】

図9において、ビデオが受信され（ステップ901）、ビデオストリームは、閾値がビデオストリームと共に送信されたかどうかを判定するため解析される（ステップ902）。閾値が送られたとき、シーンカットが出現したかどうかを判定するため、二つのフレーム間の差が閾値と比較される（ステップ903）。ビデオストリーム中に閾値が存在しない場合、ビデオストリームは、ビデオのカテゴリ（アクション、ドラマなど）がビデオストリームと共に送信されたかどうかを判定するため解析される（ステップ904）。カテゴリが送信されたとき、このカテゴリに対応した閾値が使用される（ステップ905）。カテゴリが送られなかった場合、カテゴリ及び／又は閾値は、受信したビデオのタイプに関するユーザーの知識に基づいて、受信機側で手動によって設定される（ステップ906）。次に、ビデオ索引が、カット又はキーフレームを検出するため手動で設定された閾値を使用して作成される。用語「カテゴリ」は、たとえば、ドラマ、スポーツ、アクションなどの映画タイプ、又は、たとえば、アクションの激しいシーン若しくは内容、或いは、静止したシーン若しくは内容などのようにビデオの一部分のタイプを定義する。

【0072】

アクション映画の索引を作成する場合、シーンカットを検出するための閾値は、静止タイプの放送の索引を作成する場合よりも高い値に設定される。

【0073】

或いは、非常に静的な映画、たとえば、ドラマの場合、静止シーンを検出するための閾値は、アクション映画で使用される閾値よりも下方に設定されるべきである。したがって、本発明は、ビデオのカテゴリに基づく閾値の調整に関する。アクションの激しいビデオの場合、シーン変化に対する閾値は高い値に設定され、一方、ニュース放送の場合に、シーン変化に対する閾値は低い値に設定される。このような閾値の変更は、アクション映画の場合に大量のビデオ索引が作成されることを防止し、ニュース放送の場合に少量のビデオ索引しか作成されないことを防止する。

【0074】

特定のビデオ、又は、ビデオの一部に対する閾値は、受信ビデオのヘッダに配置され得る。或いは、閾値は、たとえば、サード・パーティ・ソースによって電子番組ガイドを用いて番組毎に、若しくは、番組の一部分毎に送られる場合がある。1本の映画の中で閾値を変更することにより、全体的にはアクションの激しくないドラマではあるが、アクションの激しいシーンを含む映画を考慮することができる。また、1本の映画の中で閾値を変更することにより、アクションの激しくないフレーム間にシーン変化が存在し、アクションの激しいフレーム間にシーン変化が存在しない場合でも、多数の激しいアクションのフレームを含み、アクションの激しくないフレームを殆ど含まないようなビデオ索引の作成が回避される。サード・パーティ・サービスは、ルックアップテーブルに閾値を収容し、到来ビデオと共に閾値を受信機へ送信すること、閾値を受信機自体に供給すること、或いは、閾値を受信機へダウンロードさせることが可能である。何れの場合でも、受信されればビデオ(901)は、閾値が動的に受信されたとき(902、1003)、又は、閾値がメモリに記憶されているとき(902、1003)、閾値と比較される。

【0075】

当業者は、上記以外の多数の方法によって、閾値を受信し、設定し、使用することができる。たとえば、ホーム・ムービー、或いは、ユーザーが内容を知っている映画の場合に、ユーザーは自分自身で閾値を設定することができる。これにより、ユーザーは、ホーム・ムービーの個人向けのビデオ索引を作成することが可能になる。

【0076】

コマーシャルについては閾値を非常に高く設定してもよいことに注意する必要がある。これにより、殆どのキーフレームはコマーシャルからは選択されない。

【0077】

ビデオ、若しくは、ビデオの一部分のカテゴリーは、映画製作者によって設定し、ビデオストリーム中で供給することも可能である。受信機は、多数の閾値を記憶し、受信したビデオのカテゴリーに基づいて閾値を選択し得る(図9を参照せよ)。

【0078】

或いは、ビデオ索引付けシステムは、映画自体から判定されたビデオの受信内容に基づいて閾値を設定することができる（図10）。映画、ビデオ、若しくは、それらの一部に対し使用される閾値を判定するため、最初に、ビデオからキーフレームを選定するため、汎用的な閾値が選択される。汎用的な閾値を使用して検出されたシーン変化（カット）の数が一定の量を超える場合（1003、1004、1005）、閾値は、妥当な数のシーン変化が検出されるまで、より高い値に設定される（1006）（図10を参照せよ）。

【0079】

図10において、ビデオが受信され（1001）、二つのフレームは、二つのフレーム間の差の数を検出するため比較される（1002）。差の数は、二つのフレーム間にカットが出現したかどうかを調べるため、汎用的な閾値と比較される（1003）。次に、検出されたカットが計数される（1004）。検出されたカットの数が妥当なサイズに決められたビデオ索引を作成するには多すぎる場合（1005）、汎用的な閾値は調整され（1006）、出現が認知されるカットの数が減少する。検出されたカットの数が（作成されるビデオ索引のサイズに依存する）妥当な範囲内に収まると、この閾値はビデオ索引を作成するため使用される（1007）。

【0080】

[キーフレームフィルタリング]

キーフレームフィルタリング方法は、反復的なフレーム及びその他の選択されたタイプのフレームを除くことにより、フレームメモリに格納されるキーフレームの数を減少させるため使用される。キーフレームフィルタリングは、重要シーン検出後に、ホストプロセッサ210内でキーフレームフィルタ器240によって行われる（図2A及び2B）。重要シーン検出処理中に、多数のキーフレーム（選択されたフレーム又は画像）が抽出される。たとえば、録画済のビデオの1時間当たり2000個のキーフレームが抽出され、ユーザーが簡単に扱うためには多すぎる。しかし、ユーザーの立場では、重要シーン検出で選択された全てのキーフレームが重要であって、ビデオの視覚的内容を伝搬するために必要で

あるとは限らない。たとえば、会話シーンの場合、話者が数回表示される可能性がある。本発明は、会話シーンの場合に、話者一人について1フレームの描写だけを許可する。

【0081】

図6A及び6Bは、キーフレームフィルタリング処理の全体図である。図6Aのステップ602乃至606に示されているように、ブロック署名がフレーム内のブロック毎に獲得される。ブロック署名700は、本例の場合、図7に示されるように、8ビットにより構成され、その中の3ビットはDC署名702を表現し、5ビットがAC署名704を表現する。ブロック内でDC値以外のほかのDCT係数は、全てAC値である。

【0082】

DC署名は、ステップ602でDC値を抽出し、ステップ604でDC値が指定された値のレンジ内、本例の場合に、-2400乃至2400に収まるかどうかを判定することによって獲得される。このレンジは、図7に示されるように、予め選択された個数の区間に分割される。本実施例の場合には、8個の区間が使用されるが、映画のカテゴリーによって設定され得る画像の粒度の大小に応じて、より多くの区間、若しくは、より少ない区間が使用される。

【0083】

各区間は、図7に示されるように、所与のマッピングが割り当てられる。各DC値はこのレンジと比較され、DC値が収まる区間に対するマッピングが戻される。必要とされるビットによって表現された値は、区間の番号に対応する。本例の場合、このレンジは8区間に分割されているので、3ビットが使用される。図7に示されるように、ブロック署名700は、最初の3ビットとしてDC署名702を含み、残りの5ビットとしてAC署名704を含む。DC署名又はAC署名に割り当てられたビット数は、映画のカテゴリーに基づいて設定され得る。

【0084】

図6Aのステップ604において、ブロックに対するAC値のレンジのより良好な表現を与えるため、図7に示されるように、DC値に最も近い5個のAC値(AC1-AC5)が抽出される。ステップ606において、5個のAC値は、

それぞれ、本例の場合には200である閾値（ACthresh）と比較される。AC値がACthreshよりも大きい場合、AC署名の対応したビット706は、1のような所定の値にセットされ、AC値がACthresh以下である場合、対応したビットは0にセットされる。ACthreshは、ビデオのカテゴリに基づいて決められ得る。

【0085】

ブロック署名700は、このようにして獲得され、ブロック署名を使用することにより、単色であるフレームのような特定の画像若しくはフレームが視覚的索引から除かれる。

【0086】

単色フレームを除く簡便な方法は、ステップ602とステップ604の間で、DC署名だけに基いて行なわれる。各DC署名704は比較され、出現した特定のDC署名のカウントが保持され（ステップ660）、すなわち、各DC署名はDC値が収まる区間を表現し、本例の場合には、8個の異なるDC署名が存在する。たとえば、90%のブロック、すなわち、1782個（ $0.9 * 330 \text{ マクロブロック} * 6 \text{ ブロック}$ ）のブロックが同じ区間に収まるとき（同じDC署名を有するとき）、画像は、ステップ662において、単色であるとみされ、フレームは、廃棄されるか、又は、フレームメモリから取り除かれる（ステップ664）。或いは、別々のカウントをブロックのタイプ毎（Cr, Cb, ...）に保持し、各別のカウントを、映画のカテゴリに基づく関連した閾値と比較してもよい。

【0087】

フレーム内の各マクロブロックのブロック毎にブロック署名（Blocksig）を獲得した後、領域が判定される。領域は、本例の場合に、二つ以上のブロックであり、各ブロックは、領域内の少なくとも1個の別のブロックと接し、領域内のほかのブロックと同じブロック署名を共有する。タイミングを減少させることが望まれるならば、領域を定義するためより多くのブロックが必要とされる。フレームの各ブロック署名は、他のブロック署名と相互に比較され、領域を定めるため計数されるが、本発明は、フレーム内に領域を定めるため、領域拡張処理

のような公知技術を使用する（ステップ608）。

【0088】

領域拡張処理中に、領域カウンタは、領域内のブロックの数（サイズ）を計数するため、領域毎に保持され、16ビットで表現される。全てのフレームが領域を見つけるために解析された後、別の公知の方法が、本例の場合には、 $x-y$ 軸基準に基づく各領域の重心又は中心を見つけるため使用される（ステップ610）。 x 座標及び y 座標の両方の座標が CX 及び CY として抽出され、それぞれ、16ビットで表現される。各領域には、領域署名 $Region(Blk_sig_r, size_r, CX_r, CY_r)$ が割り当てられ、ここで、 r は領域番号を表す。領域のためのブロック署名は、領域拡張処理で決められたような最も支配的なブロック署名に基づいて決定される。

【0089】

特定の規準に基づいて、本実施例では、サイズを増大し、領域が分類され、領域フィルタリングが実行される（ステップ612）。本例の場合、大きいほうから三つの領域を除く全ての領域が除かれる。残りの3個の領域は、フレームを表現するフレーム署名に組み込まれる。フレーム署名は、本例の場合に、168ビットにより表現され、 $(Region1, Region2, Region3)$ の形式からなり、より詳細には、 $(Blk_sig1, size1, CX1, CY1, Blk_sig2, size2, CX2, CY2, Blk_sig3, size3, CX3, CY3)$ の形式である。フレーム署名のビット数は、ビデオのカテゴリーによって設定され、ビデオのカテゴリーに基づいて領域の数を変更させる。

【0090】

図6Bに示されるように、フレーム比較処理は、現在フレーム（ $F2$ ）と前フレーム（ $F1$ ）を、対応したフレーム署名に基づいて比較する（ステップ616）。本実施例の場合に、対応した $Region1$ が比較され、対応した $Region2$ が比較され、最後に、対応した $Region3$ が比較される。特に、対応した領域のブロック署名は、次式にしたがって比較される。

【0091】

$$FDiff = ABS(size1_{F1} - size1_{F2}) + ABS(size2_{F1} - size2_{F2}) + ABS(size3_{F1} - size3_{F2})$$

$FDiff < 10$ であるならば、フレームは類似しているとみなされ、更なる物体の処理が実行される（ステップ620）。

【0092】

$FDiff \geq 10$ であるならば、フレームは相異するとみなされ、どちらのフレームもこの処理の下では除去されない。明らかに、 $FDiff$ と比較する値は、ビデオのカテゴリーに応じて設定され得る。

【0093】

領域は、一般的に、「対象」を表現する。「対象」とは、物体、人、事物などを表す。対象処理は、前フレーム中の領域に表示された対象が同じ場所に在るか、或いは、非常に接近した近傍に在るかどうかを判定する。本例の場合に、対象は、フレームのプライマリ・フォーカス、又は、多数のプライマリ・フォーカスのうちの一つである。たとえば、ビデオは、ボールがフレーム内を動き回るとき、ボール遊びをしている子供に焦点（フォーカス）を合わせる。ユーザーが、画像内で対象がどこにあるかを気にしなくてもよいように、視覚的索引におけるフレーム数に制限を加えたい場合（ステップ622）、この際に、現在フレームF2がフレームメモリから除かれる（ステップ624）。

【0094】

ユーザーが、フレーム内の対象の場所に注意を払い、同じ場所若しくは非常に接近した場所に表示された対象を有するフレームだけを取り除きたい場合、幾つかの方法が使用される（対象フィルタ、ステップ626）。

【0095】

第1の方法は、以下に説明するように、ユークリッド距離を判定することによって中心を比較する。具体的には、

$$Edist = \sqrt{(CX1_{F1} - CX1_{F2})^2 + (CY1_{F1} - CY1_{F2})^2} + \sqrt{(CX2_{F1} - CX2_{F2})^2 + (CY2_{F1} - CY2_{F2})^2} + \sqrt{(CX3_{F1} - CX3_{F2})^2 + (CY3_{F1} - CY3_{F2})^2}$$

と表される。

【0096】

$E d i s t > 3$ である場合、対象は移動していると考えられ、フィルタリングは行なわれない。 $E d i s t \leq 3$ である場合、対象は略同じ場所に留まっているとみなされ、現在フレームは除かれる。 $E d i s t$ は、ビデオのカテゴリに基づく数と比較され得る。

【0097】

対象フィルタリングの第2の方法は、マクロブロックを用いてフレームを比較する。特に、対応したマクロブロック内の対応したブロックのブロック署名が比較される。たとえば、現在フレーム (F2) のMB1, 1 (フレームの位置1, 1におけるマクロブロック) のY1ブロックのブロック署名は、前フレーム (F1) のMB1, 1のY1ブロックのブロック署名と比較される。

【0098】

第一に、Y1ブロックのDC署名が比較される。DC署名が一致する場合、AC署名がビット毎に比較される。カウント (ACcount) が保持され、予め選択されたビット数、本例では5ビットのうちの4ビットが一致した場合、ブロックカウンタ (BlkCTR) がインクリメントされる。DC署名が一致しない場合、或いは、 $ACcount < 4$ である場合、次のブロックが解析される。ACcountは、ビデオのカテゴリに基づく数と比較され得る。

【0099】

マクロブロックの各ブロック (4:2:0フォーマットを使用する本例では6個のブロック) が解析される。マクロブロック内の全ブロックが解析されたとき、ブロックカウンタが検査される。BlkCTRが4である場合、ブロックは、類似しているとみなされ、マクロブロックカウンタ (MBCTR) がインクリメントされる。BlkCTRは、ビデオのカテゴリに基づく数と比較され得る。

【0100】

画像内の全マクロブロックが解析された後、MBCTRが検査される。本例の場合に、MBCTRがフレーム内のマクロブロックの75% (247、すなわち、 $0.75 * 330$ マクロブロック) 以上である場合、フレームは類似している

とみなされ、現在フレーム（F 2）はフレームメモリから除かれる。MBCTRが75%未満である場合、この時点でフレームは除かれない。MBCTRは、ビデオのカテゴリに基づく割合と比較され得る。

【0101】

単色フレームを除去する更なる方法は、領域サイズが判定されたときに行なわれる。領域サイズがフレームブロックの90%以上、すなわち、1782ブロック以上である場合、フレームは単色であるとみなされ、フレームメモリから除かれる。このフィルタリングは、上述の単色フレームフィルタリング方法よりも多数の処理を必要とする。この割合90%は、ビデオのカテゴリに基づいて変更可能である。

【0102】

キーフレーム署名に基づいて、キーフレームは、ユーザーによって望まれる可能性の最も高いキーフレームだけが維持されるように除去される。種々の閾値を使用することにより、除去されるキーフレームの数は増減し得る。

【0103】

キーフレームフィルタリング処理において、ソースビデオ中のコマーシャルの存在は、一般的に判定可能である。本発明は、ユーザーが、コマーシャルからのキーフレームを視覚的索引の一部として収容するか、或いは、除外するかどうかを選択することを可能にする。

【0104】

コマーシャルの存在は、一般的に、時間単位当たりに非常に多数のカットが含まれることによって示される。コマーシャルをより高い精度で分離すべく、ソースビデオ中のキーフレームの全体的分布を解析し、コマーシャルを伴うセグメントの頻度及び尤度の推定を試みる。

【0105】

コマーシャルは、典型的に、テレビジョン番組中の固定区間、たとえば、30分間の連続放送コメディの5乃至10分毎に分布する。コマーシャルは、1分当たりに非常に多数のキーフレームが出現するときを判定することによって分離される。特に、各キーフレームの相対時間は、他のキーフレームと比較される。

1分当たりに相対的に多数のキーフレームが出現する場合、コマーシャル、若しくは、激しいアクションのシーケンスから非常に多数のキーフレームを取得することを回避するため、この領域に対する閾値は高い値に設定する必要がある。

【0106】

重要シーケンス検出処理中に、フレームがキーフレームとしてフレームメモリに保存されたとき、関連したフレーム番号が時間コード若しくはタイムスタンプに変換され、たとえば、フレームの相対的な出現時間を示す。全てのキーフレームが抽出された後、キーフレーム密度が最後の1分間に対し計算される。

【0107】

$L1 = \text{最後の1分間のキーフレーム密度} = \text{最後の1分間のキーフレーム数} / 1800$

また、キーフレーム密度は、最後の5分間についても計算される。

【0108】

$L5 = \text{最後の5分間のキーフレーム密度} = \text{最後の5分間のキーフレーム数} / 9000$

本例では定数を3として、 $L1 > (L2 * \text{定数})$ である場合、コマーシャル中断の可能性が示される。最後に示されたコマーシャル中断の最後のキーフレームのタイムスタンプが5分よりも大きい場合、現在のコマーシャル中断が示され、最後の1分間の全てのキーフレームがフレームメモリから除去される。フィルタリング技術は、コマーシャルのキーフレームの取得を回避するため、この時間期間中に、閾値を高い値に設定するだけで実現される。

【0109】

[ビデオ検索]

ビデオテープ又はファイルが視覚的索引を有するとき、ユーザーは視覚的索引にアクセスしようとする場合がある。ビデオ検索処理は、視覚的索引を使用可能な形式でユーザーへ表示する。ユーザーは、視覚的索引を閲覧し通過し、ソーステープ又はMPEGファイルの選択されたポイントへ高速送りすることができる。図8には、検索処理が詳細に示されている。

【0110】

ステップ802において、ソースビデオが、必要に応じて、たとえば、VCR若しくは再生装置によって、視覚的索引の場所、本例の場合には、テープの先頭へ巻戻される。ソースビデオがMPEGファイル若しくはディスクである場合、ポインタが記憶場所の先頭を指示し、巻戻しを行なう必要がない。同様に、他の記憶手段は、視覚的索引の先頭へ適切に設定される。

【0111】

ステップ804において、視覚的索引は、VCRヘッド、コンピュータ、若しくは、その他のハードウェア装置によってソースビデオから読み取られ、任意のタイプの記憶装置からなる索引メモリに保存される。ステップ806において、たとえば、VCR内のプロセッサは、ソースビデオの索引メモリからキーフレームを取り出す。ステップ806において、取り出されたキーフレームは、サイズを縮小するため、本例の場合には、たとえば、120×80画素に縮小するため処理される。勿論、これ以外のフレームサイズが自動的に、若しくは、ユーザーによって手動的に簡単に選択される。

【0112】

処理されたフレームは、次に、ステップ808においてホストプロセッサ210へ転送され、処理されたキーフレームを表示メモリに書き込み、コンピュータディスプレイ、テレビジョン画面などのユーザー・インタフェースに処理されたキーフレームを表示する。

【0113】

ステップ810において、ビデオ索引が読み取られた後、ソースビデオは停止される。ビデオ索引付けシステム、又は、ソフトウェアは、ステップ812において、キーフレームを、コンピュータモニター、若しくは、テレビジョン画面のようなディスプレイに表示させることができる。ステップ814において、視覚的索引は、ユーザーが希望する場合に印刷される。ユーザーは、ステップ816において、視覚的索引から特定のキーフレームを選択する。ユーザーが特定のキーフレームにおけるソースビデオを視聴したい場合、ソーステープは、キーフレームが抽出された場所から、ソーステープ上の対応したポイントへ自動的に早送りされ、これにより、ソーステープを再生することができる（ステップ818）

。或いは、カウンタが表示され、ユーザーは、視覚的索引から選択されたキーフレームまで、ソースビデオを高速送りするか、ソースビデオを再生することができ

【0114】

本発明は、重要シーン検出処理を省き、キーフレームフィルタリング処理だけを実行してもよい。しかし、現在広く普及しているプロセッサを使用する処理は非常に遅くなる。

【0115】

本発明の更なる特徴は、ユーザーが任意のポイントでビデオテープの再生を停止し、そのビデオテープの視覚的索引にアクセスできることである。このためには、ビデオテープがセッション中に最初に使用されるときに、視覚的索引を記憶するメモリ若しくはバッファが必要になる。

【0116】

本発明は、DCT係数を使用する例に関して説明されているが、波長係数のような代表値、又は、画像の部分領域に作用し画像の部分領域に対する代表値を与える関数を代わりに使用してもよい。これは、重要シーン検出と、キーフレームフィルタリングの両方で使用される。

【0117】

請求項に記載されているように、本発明によるビデオシステムは、ビデオのフレームを受信する受信機と、ビデオの特定の 카테고리 に対する閾値を受け取る閾値メモリと、第1のフレームと後続のフレームの差を判定するため第1のフレームを後続のフレームと比較し、この差を閾値と比較し、差が閾値よりも大きい場合に両方のフレームの間でシーン変化が生じているとみなし、差が閾値よりも小さい場合にシーン変化が生じていないとみなす比較器と、を含む。

【0118】

選択的に、差が複数のフレームに対する閾値よりも小さい場合、シーンは静止シーンであるとみなされる。シーンが静止シーンであるとみなされるかどうかと

は無関係に、閾値は到来ビデオから受け取られ、或いは、閾値はサード・パーティ・ソースから受け取られる。本発明のシステムがこのように構成されている場合、更に、閾値設定装置が設けられ、閾値設定装置は、

ビデオのカテゴリを検出するカテゴリ検出器と、

検出されたカテゴリに基づいて閾値を設定する閾値設定器と、

を含む。閾値は、受信機側で設定してもよい。

【0119】

上述の通り、本発明は、好ましい実施例に関して説明されているが、上記の本発明の原理に含まれる本発明の実施例の変形は、当業者にとって明らかであり、本発明は、上記の好ましい実施例に限定されることはなく、このような実施例の変形を包含することが意図されている。

【図面の簡単な説明】

【図1】

ビデオ保管処理の説明図である。

【図2A】

視覚的索引を作成する際に使用される装置の概要図である。

【図2B】

視覚的索引を作成する際に使用される装置の概要図である。

【図3】

フレーム、マクロブロック及び数個のブロックの説明図である。

【図4】

ブロックの数個のDCT係数の説明図である。

【図5】

マクロブロック及び数個のブロックをDCT係数と共に示す図である。

【図6A】

キーフレームフィルタリング処理の手順の説明図である。

【図6B】

キーフレームフィルタリング処理の手順の説明図である。

【図7】

マクロブロックと、マクロブロックのブロック署名との関係の説明図である。

【図 8】

ビデオ検索処理の説明図である。

【図 8 A】

ビデオキーフレームシステムの構成図である。

【図 9】

閾値検出装置の説明図である。

【図 10】

閾値設定装置の説明図である。

【図1】

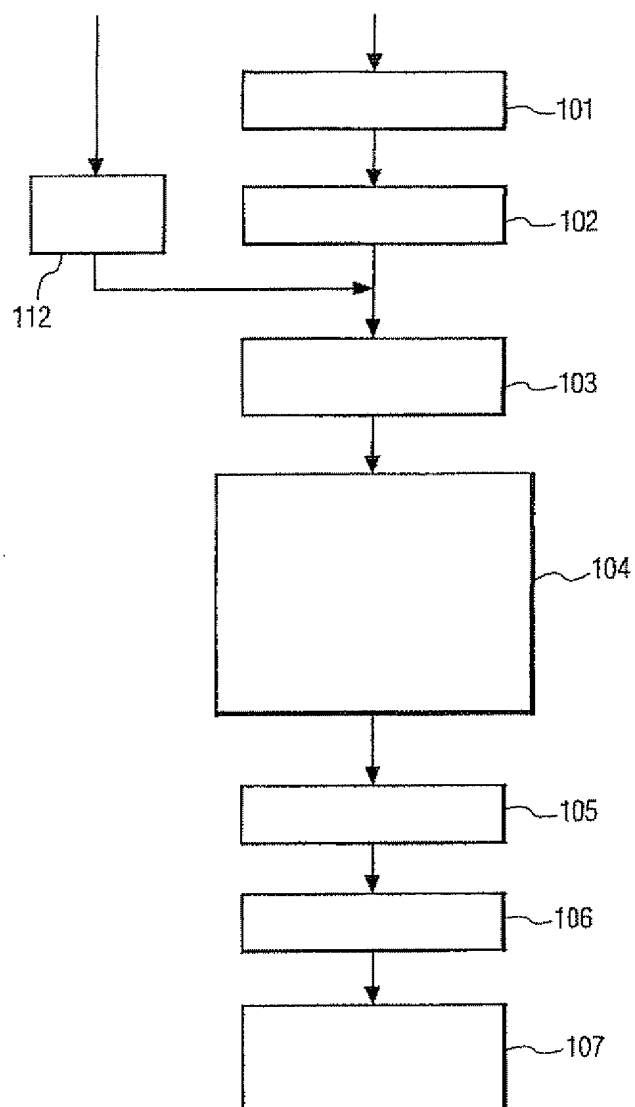


FIG. 1

【図2A】

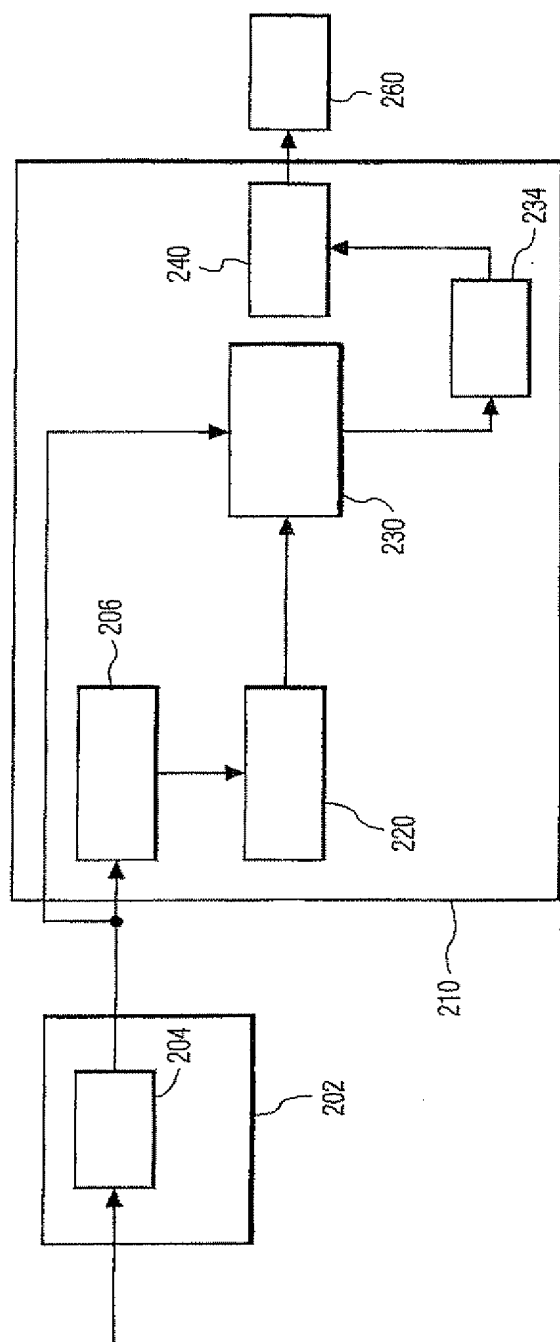
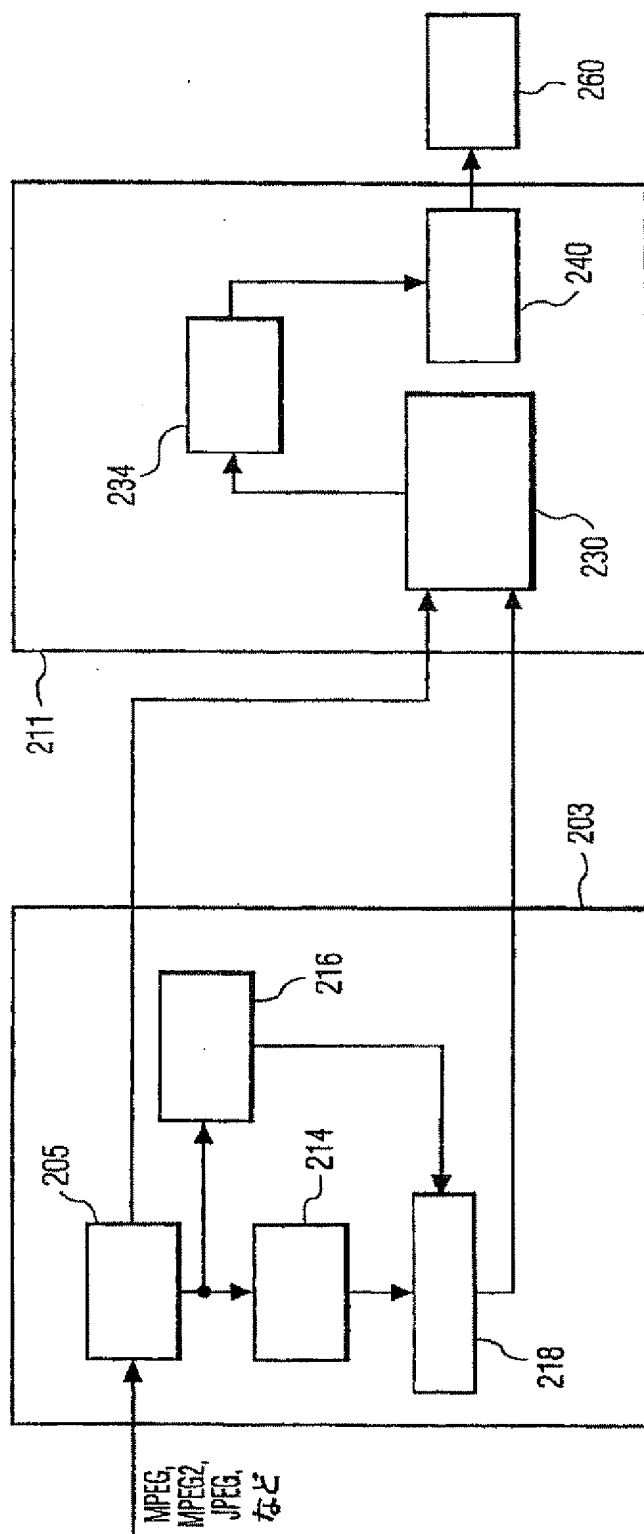


FIG. 2A

【図2B】



【図3】

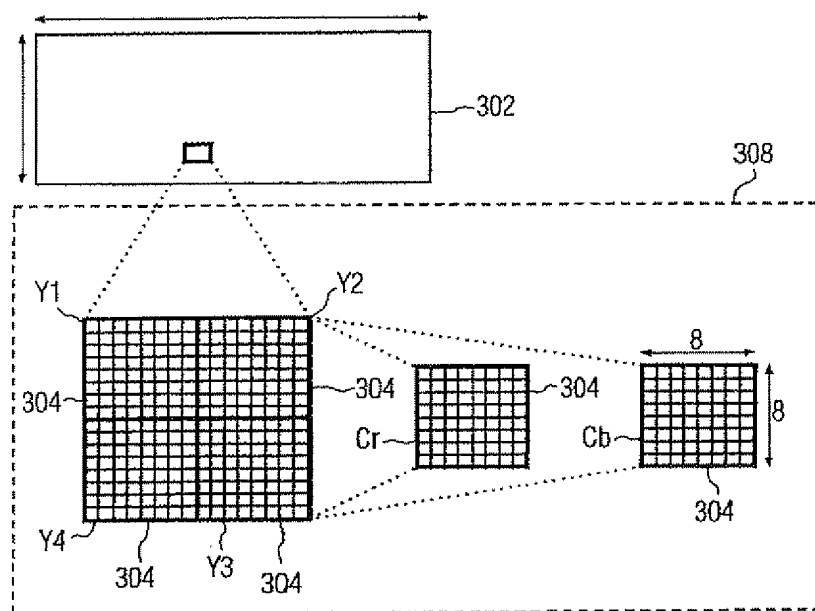


FIG. 3

【図4】

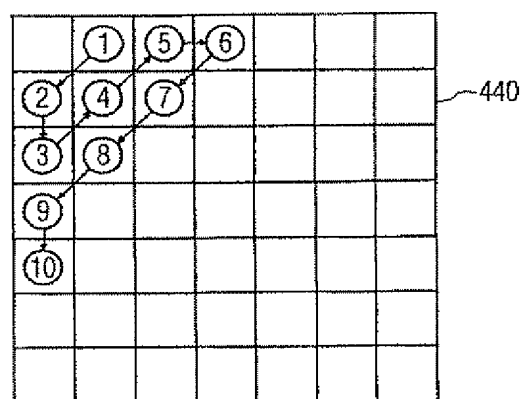
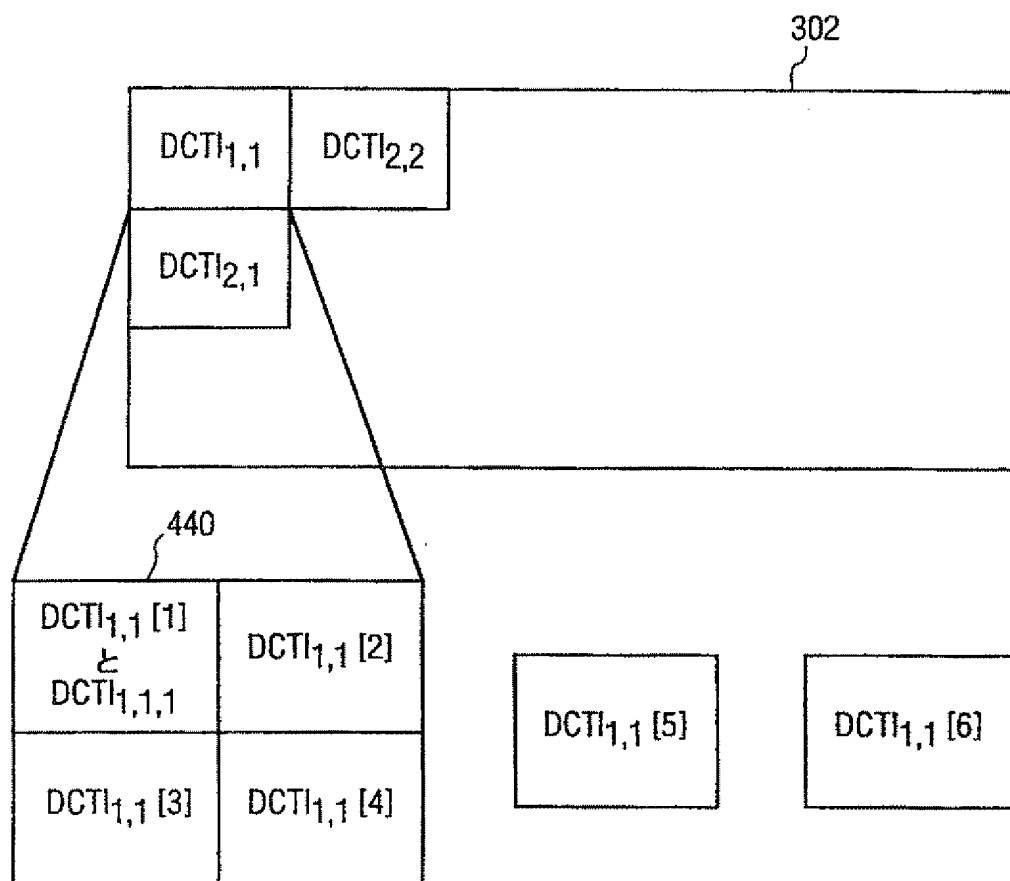


FIG. 4

【図5】



【図 6 A】

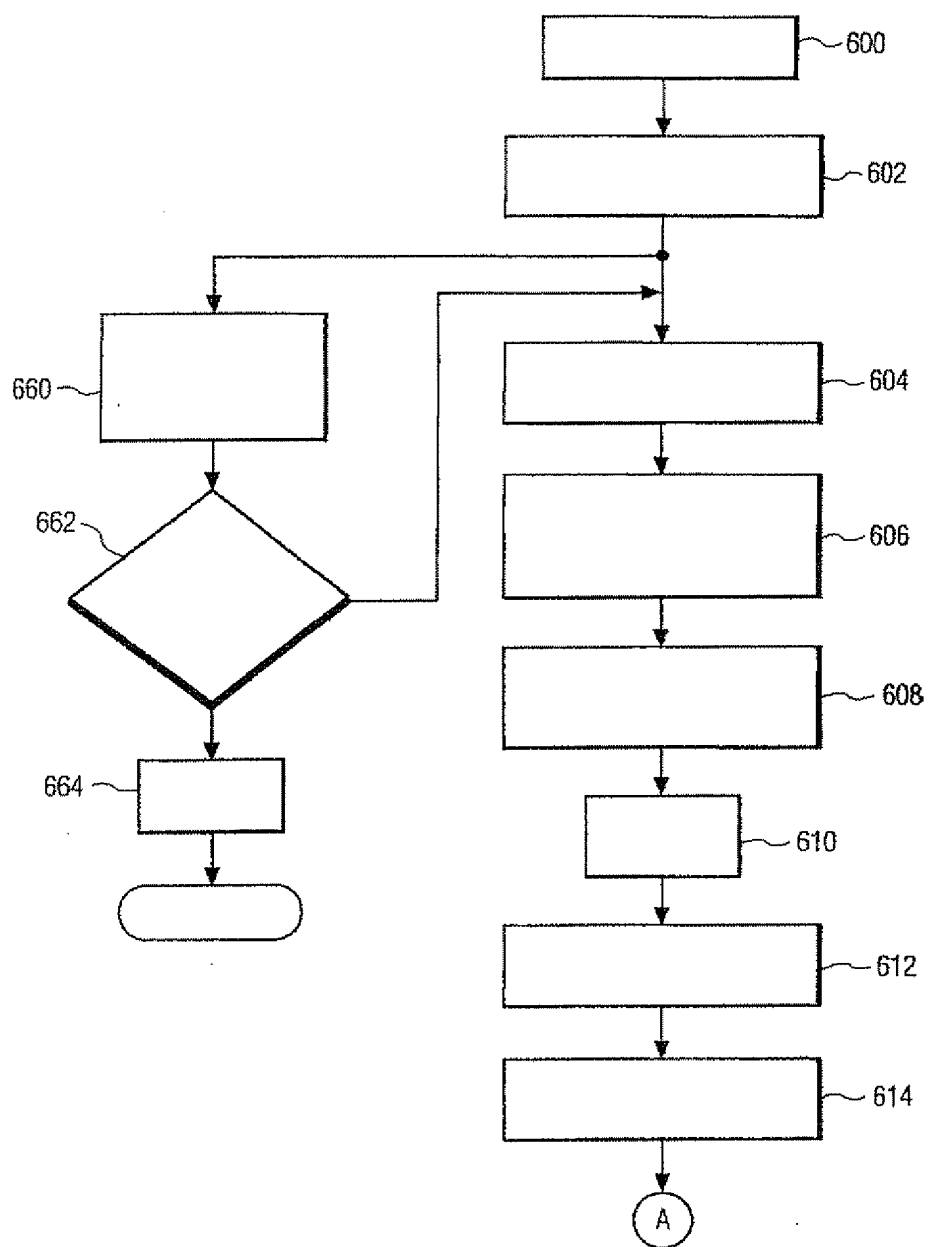


FIG. 6A

【図 6 B】

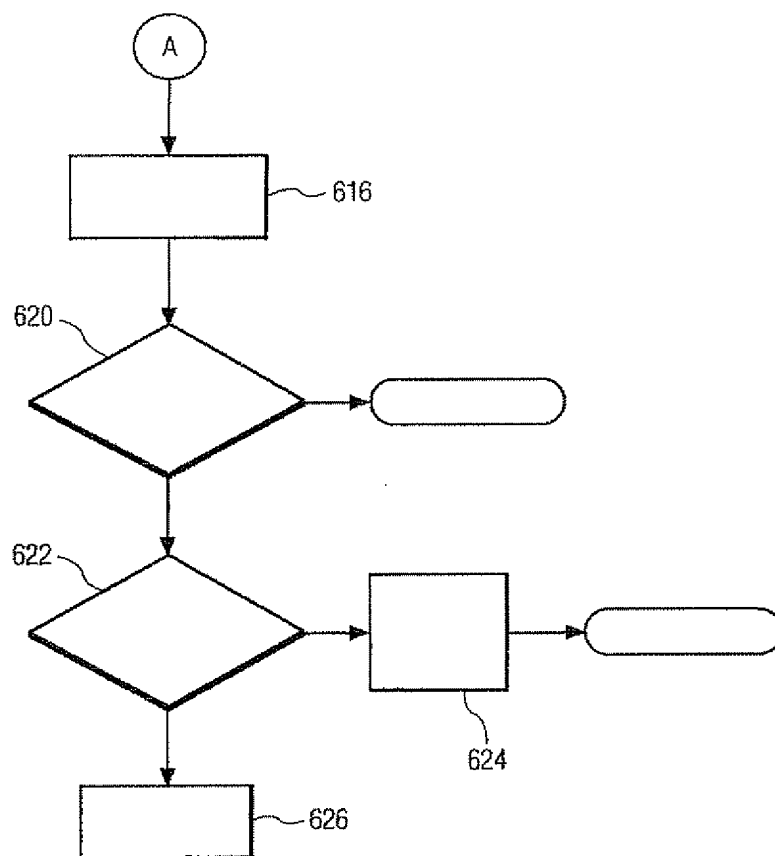
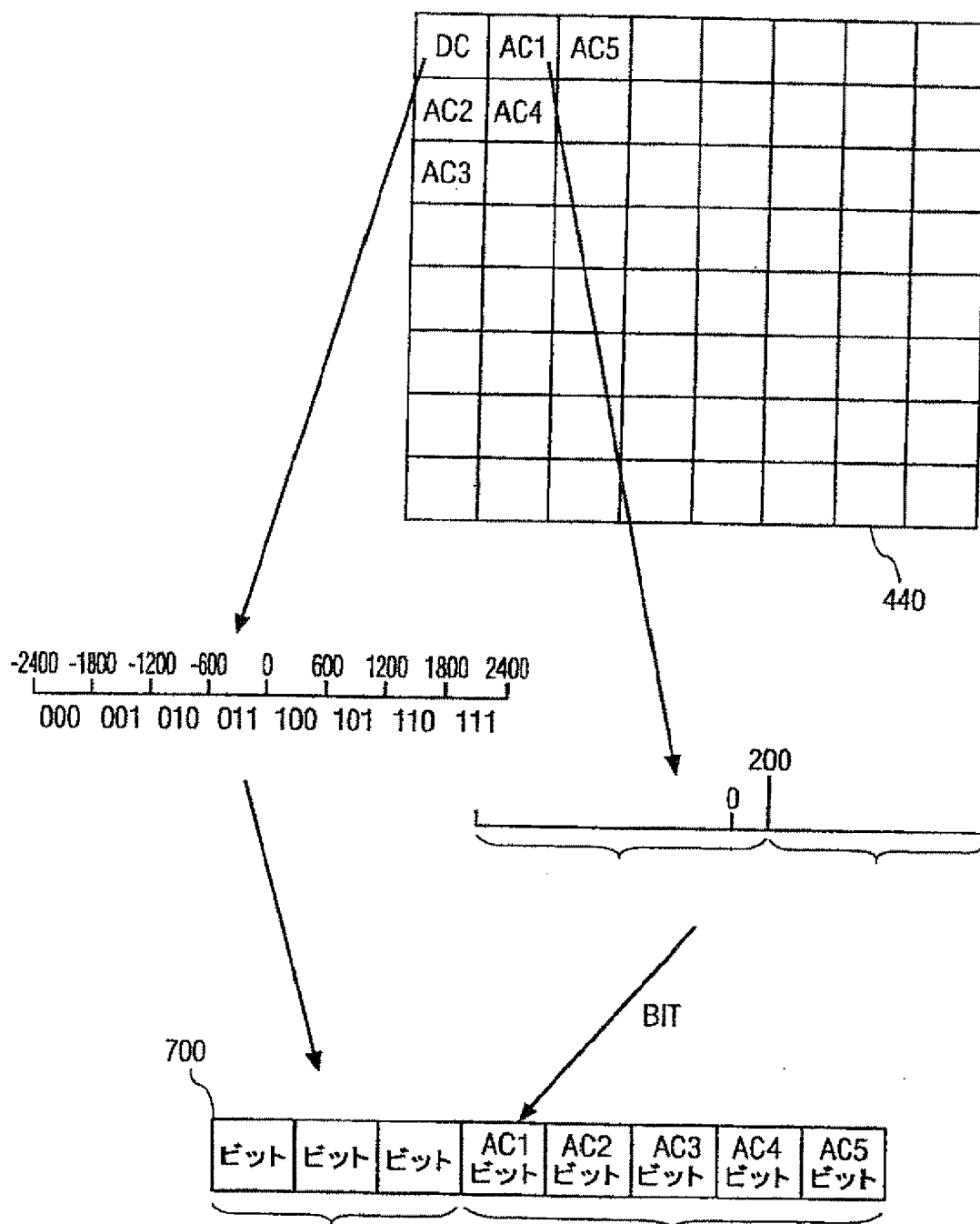


FIG. 6B

【図7】



【図 8】

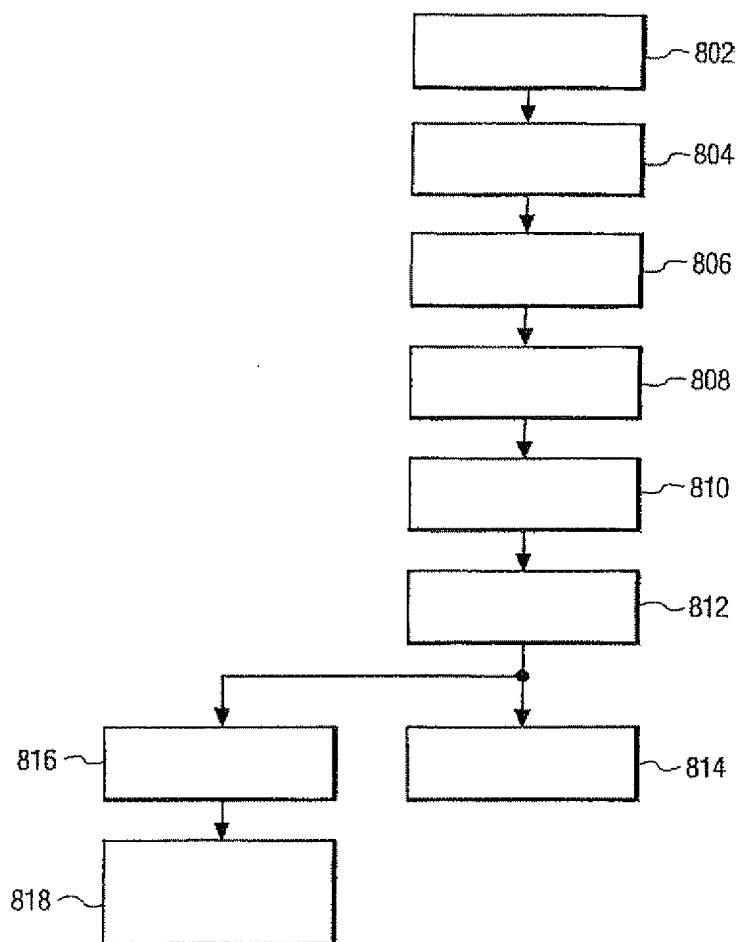


FIG. 8

【図 8 A】

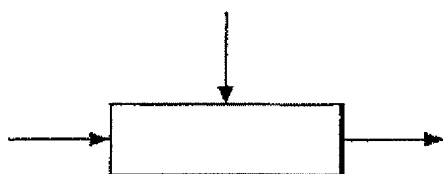


FIG. 8A

【図9】

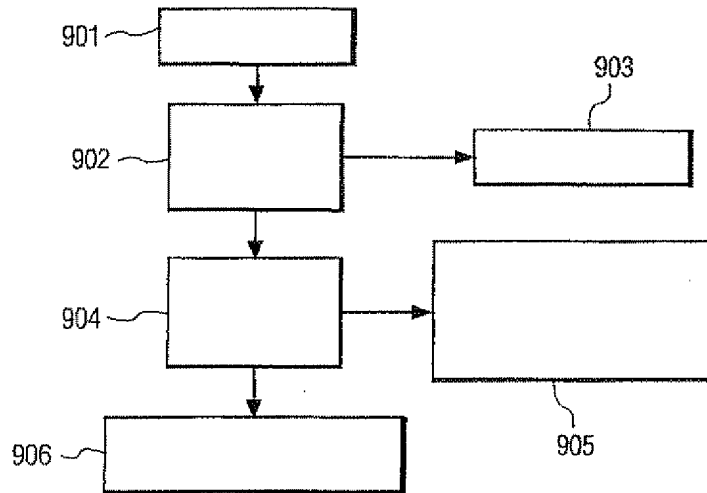


FIG. 9

【図10】

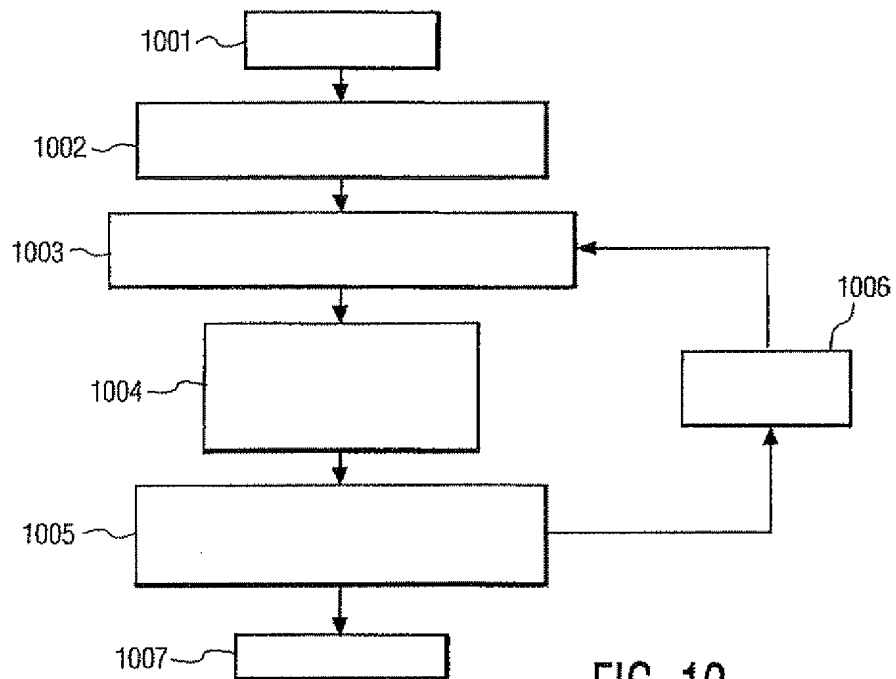


FIG. 10

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/EP 00/10161
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 H04N7/36 H04N5/14 G06F17/30		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 G06F H04N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, PAJ		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 55943 A (KONINKL PHILIPS ELECTRONICS NV ; PHILIPS SVENSKA AB (SE)) 10 December 1998 (1998-12-10) page 1, line 26 -page 18, line 24 claims 1-3,5,6,8-11	11,12,15
Y		1-4,13, 14
A	---	5-10
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another priority or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "Z" document member of the same patent family		
Date of the actual completion of the international search 5 February 2001		Date of mailing of the international search report 13/02/2001
Name and mailing address of the ISA European Patent Office, P.O. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040. Tx. 31 651 epo nl. Fax: (+31-70) 340-3016		Authorized officer Fassnacht, C

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int. l. Application No.
PCT/EP 00/10161

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category ^a	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	LIENHART R ET AL: "On the detection and recognition of television commercials" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON MULTIMEDIA COMPUTING AND SYSTEMS. LOS ALAMITOS, CA, US, 3 June 1997 (1997-06-03), pages 509-516, XP002154465 page 511, right-hand column, paragraph 3 page 515, left-hand column, paragraphs 3,4	1-4, 13, 14
A	-----	5-12, 15
A	DIMITROVA N ET AL: "VIDEO KEYFRAME EXTRACTION AND FILTERING: A KEYFRAME IS NOT A KEYFRAME TO EVERYONE" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON INFORMATION AND KNOWLEDGE MANAGEMENT. CIKM, US, NEW YORK, ACM, vol. CONF. 6, 10 November 1997 (1997-11-10), pages 113-120, XP000775302 ISBN: 0-89791-970-X the whole document	1-15
A	-----	1-15
	MC GEE T ET AL: "PARSING TV PROGRAMS FOR IDENTIFICATION AND REMOVAL OF NON-STORY SEGMENTS" PROCEEDINGS OF THE SPIE, 1998, XP000972571 the whole document	

INTERNATIONAL SEARCH REPORT

information on patent family members

Int. Patent Application No.
PCT/EP 00/10161

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9855943 A	10-12-1998	US 6137544 A	24-10-2000
		EP 0944874 A	29-09-1999

フロントページの続き

(72)発明者 ディミトロワ, ネヴェンカ
オランダ国, 5656 アーアー アインドー
フェン, プロフ・ホルストラーン 6

(72)発明者 エレンバース, ジャン エイチ
オランダ国, 5656 アーアー アインドー
フェン, プロフ・ホルストラーン 6

F ターム(参考) 5C053 FA20 GB21 HA29 JA21 KA03
KA11 KA22 KA24 LA07
5C059 MA23 MC32 PP04 SS11 TA68
TB07 TB08 TC00 UA11 UA38
UA39

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
10 May 2001 (10.05.2001)

PCT

(10) International Publication Number
WO 01/33863 A1

(51) International Patent Classification⁷: **H04N 7/36,**
5/14, G06F 17/30

Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).
ELENBAAS, Jan, H.; Prof. Holstlaan 6, NL-5656 AA
Eindhoven (NL).

(21) International Application Number: PCT/EP00/10161

(22) International Filing Date: 13 October 2000 (13.10.2000)

(74) Agent: **WHITE, Andrew, G.**; Internationaal Octrooibu-
reau B.V., Prof. Holstlaan 6, NL-5656 AA Eindhoven
(NL).

(25) Filing Language: English

(81) Designated States (*national*): JP, KR.

(26) Publication Language: English

(30) Priority Data:
09/433,258 4 November 1999 (04.11.1999) US

(84) Designated States (*regional*): European patent (AT, BE,
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,
NL, PT, SE).

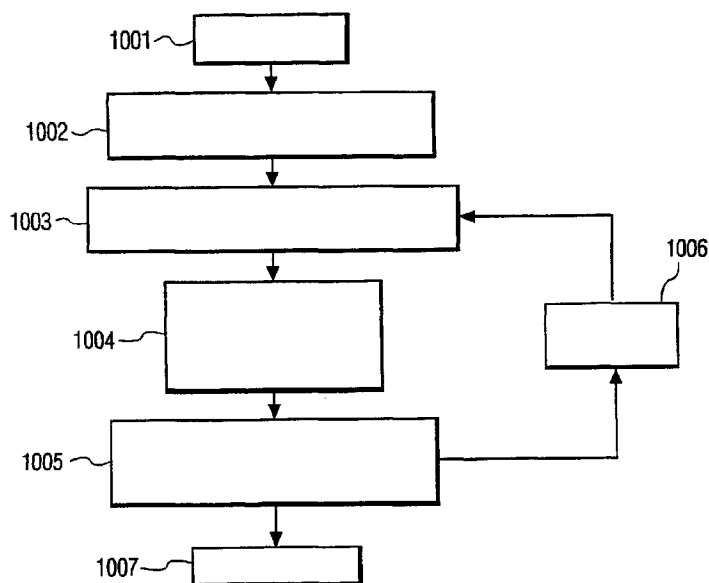
(71) Applicant: **KONINKLIJKE PHILIPS ELECTRON-
ICS N.V.** [NL/NL]; Groenewoudseweg 1, NL-5621 BA
Eindhoven (NL).

Published:
— *With international search report.*

(72) Inventors: **MC GEE, Thomas**; Prof. Holstlaan 6,
NL-5656 AA Eindhoven (NL). **DIMITROVA, Nevenka**;

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

(54) Title: SIGNIFICANT SCENE DETECTION AND FRAME FILTERING FOR A VISUAL INDEXING SYSTEM USING DYNAMIC THRESHOLD



(57) Abstract: A video indexing system analyzes contents of source video and develops a visual table of contents using selected images. A system for detecting significant scenes detects video cuts from one scene to another, and static scenes based on DCT coefficients and macroblocks. A keyframe filtering process filters out less desired frames including, for example, unicolor frames, or those frames having a same object as a primary focus or one primary focuses. Commercials may also be detected and frames of commercials eliminated. The significant scenes and static scenes are detected based on a threshold which is set based on the category of the video.